

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGGGTGGT
TTCAGCAAGGCCTCAGTTCTTCAGCCCTGTAATTGGACATCTGCTGCTTCATATT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTATGCATTG
CTACCATTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTGTACTTGAATACTGAGTTGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTCAGACCATCCTTCACCAAATGCAGCCAAAATCCATGGCAAACAAAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTCCTCTT
GGTTTTCTGACTTACATTGTAATTTCAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCTCTATGACACTGCACCTGCCCTATTAACAATGAACGAACACGGCTACTTCCA
GAGATATTGATGAAAGGATAAAATATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTCAACCACCTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTGATAGATTCTAAAGGATATCAT
CAAGAAGACTATTAAAACACCTATGCCTATACTTTTATCTCAGAAAATAAGTCAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMQPKIHGKQFWIRLLLVIWCGVSALSMLTCSVLHSGNFGTDLHQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACCGTGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCTCGGGAGTGGAAAGTGGAGGCAGGAGCCTC
CTTACACTTCGCC**CATGAG**TTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTG
GATTGGGTGGCTTTCTTCATGCGCCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTGCATTTCTGCACCCTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTCACTGGAAAATGAACCTGTGTAACTCTGCTGATCCTGG
TTTCATGGTGCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACACGA
CTGCTTTCTGTCTTATGGCTGACCTTATGTATTTCTTGAAACTAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTGGATTGGTGCTGTCAACTGCCATACACTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCACTGCTGCAAACCATGGATATGAT
CATAGAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGCTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAACAGAACACTCCAAAACCTCAAGGGAAATATTTAATTTCTTGGTT
ACTTTCTCTATTACTGTGTTGGAAAATTTCATGGCTACCATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCCCTCATTCTGGAAATAATCGTCACATCCATCAGGGAT
TGCTGATCACTCTTACCAAGTTCTTATGCCATCTAGCAGTAAGTCCCTCAATGTCATTGCTCG
CTATTAGCACAGATAATGGCATGTACTTGTCTCCTCTGTGCTGATCCGAATGAGTATGCCTT
AGAATACCGCACCATAACTGAAGTCCTGGAGAACACTGCAGTTCAACTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCAC**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTAA
GATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTATAAACAAACAAATGCTATGGTAGC
ATTTTCACCTCATAGCATACTCCTCCCCGTCAAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAACACTCAAGACAATACTCAGCAGAGGCATCCGTGTTGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCCAAGAAACTAAAGGTGAAAATACACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTAGGATTCCGTTAAGGTTCACATGGAAAAGGTTAGCTT
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCATGCCCAACTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSSRYFHWKMNLCVILLILVFVMPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDP
FPILSPKHGILSIEQLISRGVIGVTLALLSGFGAVNCPTYMSYFLRNVTDILALERRLLQ
TMDMIISKKRMAMARRTMFQKGEVHNKPSGFWMGIKSVTTSASGSENLTLIQQEVDALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSMPLERYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGAGGCCAGGTGCTTCAGCCTGGTGTGCTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCTTGTGAGCAAAAGGCAGACCAGCAGCTGAATTACAG
AAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTGGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTGTCAGCTCTAG
GATTAGCCAAACCCAAAGTGTGGAAAAATGGGGTGGGTGCTGATTGGAAGGTTCCAGTGA
GCCGACAGTTGCAGCCATTGTTACAACACTCATCTGATACTTGGACTAACTCGTGCATCCAGAA
ATTATCACCACCAAAGATCCCATTCAACACTCAAACACTGCAACACAAACAAACAGAATTATTGT
CACTGACAGTACCTACTCGGTGGCATCCCTACTCTACAATACCTGCCCTACTACTCCTC
CTGCTCCAGCTCCACTTCTATTCCACGGAGAAAAAAATTGATTGTCAGAGTGGTTATG
GAAACTAGCACCAGTCTACAGAAACTGAACCATTGTTGAAAATAAGCAGCATTCAAGAATGA
AGCTGCTGGTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTGCTCTCCTCTTGGTGTG
CAGCTGGTCTGGATTGCTATGTCAAAAGGTATGTGAAGGCCTCCCTTACAAACAAGAAT
CAGCAGAAGGAATGATGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC
GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCCTTACCCCTGCCAGCTGGGAAATCAAAGGCCAAAGAACCAAAGAAGAAAGTCCA
CCCTGGTTCTTACTGGAAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT
GCCCTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCTACCTCCAAAGCTCCACGGCCT
TTCTAGCCTGGCTATGCTCTAATAATATCCACTGGAGAAAGGAGTTTGCAAAGTGCAAGGAC
CTAAAACATCTCATCAGTATCCAGTGGAAAAAGGCCTCTGGCTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGGACTGAGACCAAGGCTTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGCTCTGAGCCGGTAAGAGCAAAAGAAT
GGCAGAAAAGTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAA
GCTAAAATAAGAAATAGAACAAAGGCTGAGGATAACGACAGTACACTGTCAGCAGGGACTGTAAAC
ACAGACAGGGTCAAAGTGTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGTCTCTACCACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAACTCTTATAAATTCTATTCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCTCAAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTTTGATATTCTAGCTTACCTCTACCTCCAAACTAAT
TTTATTGCTGAGACTAATCTTATTCTAATATGGCAACCATTATAACCTTAATT
TATTATAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAAGCACATTAAAAGTGCC
ATTAACAAATGTACTAGCCCTCTTCCAACAAGAAGGGACTGAGAGATGCAGAAATT
TGTGACAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVSPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAASFNEAGFGGVTALLVLALLFGAAAGLGFCYVK
RYVKAFTPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDFNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCCTCCGCACCGCGGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCGGC
GGCCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCGGGGCGG
CGGCTGCGGGCGCAGACGGAGATGCAGCGGTTGGGCCACCCCTGCTGTGCTGCTGCTGGCG
CGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAGGCCGGCCCG
GCTCTCAGCTACCCGCAGGAGGAGGCCACCCCTCAATGAGATGTTCCCGAGGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTACACATCTGTGGGAGACGAAGAAGGCAGAAGGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGAGGATGCTCTGCACCCGGACAGTGAGTGTGGAGACCA
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGACAGTGAGTGTGGAGACCA
GCTGTGTCTGGGTCACTGCACCAAAATGCCACCAGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGCTGTGCTGTGCCCTCAGAGAGGCCGCTGTTCCCTGTG
ACACCCCTGCCGTGGAGGGCGAGCTTGCATGACCCGCCAGCCGGCTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTCGTGGGAGGCCGTGACCAAGATGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGTGGGGAGCCTGCCAGC
CCGCCGCTGCACTGCTGGAGGGAAAGAGATTTAGTCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCCCCAGGTGTGCTTAGCGTGGCTGACCAGGCTTCTTCCTA
CATCTTCTCCCAGTAAGTTCCCCTGGCTTGACAGCATGAGGTGTTGTGCATTGTTAGCT
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGGCTTGGAGAGTCAGGCAGGGTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATTGGCTGCTTGCCCTTACCAAGTGGCAGACAGCCG
TTTGGTCTACATGGCTTGATAATTGTTGAGGGGAGGAGATGAAACAATGTGGAGTCTCCCTC
TGATTGGTTGGGAAATGTGAGAAGAGTGCCTGCTTGCAAACATCAACCTGGAAAAATG
CAACAAATGAATTTCACGCAGTTCTCATGGCATAGGTAAAGCTGTGCCCTCAGCTGTTGC
AGATGAAATGTTCTGTTCACCCCTGCATTACATGTGTTATTCAGCATGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGGG
AGGGGGTATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCCACACCAGCCTGGTGCACCAAAAGTGCCTCCAAAAGGAAGGAGAATGGGAT
TTTCTTGAGGCATGCACATCTGAAATTAGGTCAAACATAATTCTCACATCCCTCTAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGCAGCCGTCTTAATGAAGACAATGAT
ATTGACACTGTCCTCTTGGCAGTTGCATTAGTAACTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAATTCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTCACTGGACTGTTGCCACCATGTATTCACTCCAGAGTTCTAAAGTTAAAGTGC
CATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPGASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGTGCCTAGGGCTGATGTGCGTCTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGGCTCTGGGGCTCTGGACCCCTT
AACTGGGTACTGGCCCTGGGCCAATGCGCCTCGCTGGAGCCTTGCCTCCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCCTTAATCTCTGCCTCATCCGACACTCCGTTACC
ACACTGGGTATTGGCATTGGAGGCCCTCATCCTGACCTGTGCAGATAGCCCGGGTCATTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTTCTAAACCGCAATGCATACTCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCAAAATGCGTTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTGGAGCTGCTGGT
GGTCGGAGGCCTGGGGCTGTCTCTCTCTCTCCGGTCGCATCCCGGGCTGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATGCCAGCGCTTCTCAGCTTCGGCATGTGTGGACACGCTCTCCTTGCTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTCGCCTACAGGT
CTCCATTTGTGGTAAAAAAAGTTTAGGCCAGGCCTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTGAGACCAGCCTGGCAACATGGT
AAACCTCCGTCTATTAACAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCC
GCTACTCGGAGGCTGAGGAGAATCGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAACAAACAAA
AAGATTATTAAGATATTTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHVNSSCPGLMCVFQGYSSKGLIQRSVFNQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLFFGKLLVGGVGVLFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTCCTCTGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCAGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGTCTGCAGGGCACATCGACTGTGGCTCCCTGCTGGCTACCGCGCTGTCTACCG
CATGTGCTCGCCACGGCGCCTCTTCTTCTTACCCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTGGTTAAGTCCTGATCCTGGTGGCCTC
ACCGTGGGTGCCCTCTACATCCCTGACGGCTCCTCACCAACATCTGGTCTACTTCGGCGTGTGGG
CTCCTCCTCTCATCCTCATCCAGCTGGTGCTCATCGACTTGCACACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCCTTCTTACTCTCCTC
TTCTACTTGCTGTCGATCGCGGCCGTGGCGCTGATGTTATGTACTACACTGAGCCCAGGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTAACCTCACCTCTGTGCTGCGTGTCCATCGCTGCTGTCCCTGC
CCAAGGTCCAGGACGCCAGCCAACTCGGGTCTGCTGCAGGCCCTGGTCATCACCCCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGATGCCCGAGCATTG
TGGGCCTCATCATCTCCTCTGTGCACCCCTTCATCAGTCTGCGCTCCTCAGACCACGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGCCTTGACAACGAGCAGGACGGCGTACCTACAGCTACTCCTCTCC
ACTTCTGCCTGGTGCTGGCTCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCCAGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCCTGCGAACCGCGACTTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCTC
CCCACACCAATGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTCCTCCCTCCCTGTTGCCACTCAGCATCTCGGATGAA
AGGGCTCCCTGTCCCTCAGGCTCACGGGAGCAGGGCTGCTGGAGAGAGCAGGGAACTCCCACCA
TGGGGCATCCGGCACTGAAGCCCTGGTGTCCCTGGTCACGTCCCCAGGGACCCCTGCCCTCCTG
GACTTCGTGCCTTACTGAGTCTAAGACTTTCTAATAAACAAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILQLVLLIDFAHSWNQRWLKGAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNLSMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGCGGCCGGCCAGGAACCACCGTTAAGGTGTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTGCGCAATATCCATTCCATCAACCCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTCTTCTGGCAGTTTCGATTAAAGTGTAAACTTGCAATGCTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAAG
TGATCCTTCGAAGCTTCTCAAGGGCTTGGCTATGTGCTGCCATCATTTCATTCA
CTTGCCCTGGATTGAGACGTGGCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGCTTGATG
GTCAGTTTATTCCCCTCTGAATCCGAAGCAGGATCTGAAGAACGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTAGAACTATGAGTACTACTTTGTTAAATGTGAAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTACGTC
CACTGCTGGCTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCACTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGATACTTTTATTCAATTAC
AGAATGGAATTTTTGTTCATGTCAGATTATTGTTAAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTTTGTCAGTTAAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTATTCTGTTGCATTATGTGATGGCCTGAAGTGTGGA
CTTGCAAAAGGGAAAGAAAGGAATTGCGAACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAACTCTTATTGAAATGC
ACAAAATGACTAAACCATTCATATCATGTTCCCTTGCCTCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTCGGGACCCAGGACCCCCCTGGGCCGACCCGCCAGGAAAGACTGAGG
CCGCAGCCTGCCCGCCGGCTCCCTGCGCCGCCGCCCTCCCGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCTGCTACTGCCCTGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAGCCACAGACAGTCTCTGCACTGCCGCCAGGGACACGGTGC
CGAGACGTGCCACCCGACACGGTGGGCTGTACGTCTTGAGAACGGCATCACCATGCTGACGC
AGGCAGCTTGCCGGCTGCCGGGCTGCAGCTGGACCTGTACAGAACAGATGCCAGCC
TGCCCAGCAGGGTCTTCCAGCCACTGCCAACCTCAGAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATACCAATGAGACCTTCCGTGCCCTGCAGACACGCTGACCGCCTCTGGAGCTAAGCTGC
CCGCATCCGCCACATCCAGCCTGGTGCCTCGACACGCTGACCGCCTCTGGAGCTAAGCTGC
AGGACAACGAGCTGCCGGACTGCCCTGCCCTGCAGCTGCCCTGCTGCTGGACCTCAGC
ACAACAGCCTGCCCTGGAGCCGCATCTGGACACTGCCAACGTGGAGGCGCTGCC
GGCTGGTCTGGGGCTGCAGCAGCTGGAGGGCTTCAGCCGCTGCGCAACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCCTCCGGGCTGACG
CGCCTGCCGCTGGCCGGCAACACCCGCATTGCCAGCTGCCGCCAGGACCTGCCGGCCTGGC
TGCCCTGCAGGAGCTGGATGTGAGAACCTAACGCTGCAGGCCCTGCCCTGGCGACCTCTGGGCC
TCTTCCCCGCCCTGCCCTGCTGGCAGCTGCCGAACCCCTAACACTGCGTGTGCCCTGAGC
TGGTTGGCCCTGGGTGCGCGAGAGCCACGTACACTGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCGCCAACAGAACGCTGCCGGCTGCTCCCTGGAGCTTGAACGCTGCCGACTTGGCTGCCAG
CCACCACCAACAGCCACAGTGCCACACAGGAGGCCACAGCCTGTCT
TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGGCCGGCACTGAGGCCACCGTCCACCTGCCTCA
CACTGCCCAACCGACTGTAGGGCTGCCCCAGGCCAGGACTGCCACCGTCCACCTGCCTCA
ATGGGGCACATGCCACCTGGGACACGGCACCACCTGGCTGCTTGTGCCCGAAGGCTTCACG
GGCCTGTACTGTGAGAGCCAGATGGGACAGGGCACGCCCTACACCAGTCACGCCGAG
GCCACCACGGTCCCTGACCTGGCATCGAGCCGGTGAGGCCACCTCCCTGCGTGGGCTGC
AGCGCTACCTCCAGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGAACCTATCG
GGCCCTGATAAGCGGCTGGTGACCGCTGCACTGCCCTCGCTCGCTGAGTACACGGTACCCCA
GCTGCCGCCAACGCCACTTACTCCGTCTGTGTCATGCCCTGGGGCCGGGTGCCGGAGG
GCGAGGAGGCCCTGCCGGGAGGCCATACACCCCAAGCCGTCCACTCCAACCACGCCAGTCACC
CAGGCCCGAGGGCAACCTGCCCTCTGCCCTGGCCGGTGCCTCTGCC
GCTGGCTGCCGTGGGGCAGCCTACTGTGTCGGGGGGCCATGGCAGCAGGGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGGCTGGGACTGGAGGGAGTGAAGGTCCCTGGAG
CCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGGGTCTGAGTGTGAGGTGCC
ACTCATGGCTTCCCAGGGCTGCCCTCAGTCACCCCTCCACGCAAAGCCCTACATTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGCTCAGCCAGTGCAGATGCCAGGCCCTCTGCTGCC
ACACCACGTAAGTTCTCAGTCCAAACCTCGGGATGTGTCAGACAGGGCTGTGACCACAGCT
GGGCCCTGTTCCCTGGACCTCGGTCTCTCATCTGTGAGATGCTGTGCCAGCTGACGCC
CTAACGTCCCAGAACCGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGCAGGCCGGCCCTGCCATGTGCTGGTAACGCATGCCCTGGCTCTGCTGGCTCTCCAC
TCCAGGCCGGACCCCTGGGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGCTTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGGAAAGGAAGATGC
TTTAGGAACATGTTGCTTTAAAATATATATTATAAGAGATCCTTCCATTATTCT
GGGAAGATGTTTCAAACTCAGAGACAAGGACTTGGTTTGTAAAGACAAACGATGATATGAA
GGCTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLLALGPGVQGCPGQCQCSQPQTVFCTARQGTTPRDVPPDTVGLYFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDTANVE
ALRLAGLGLQQLDEGLFSRLRNLHLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLPGDLSGLFPRLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTLGIEPVSPSLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAQDKGQVPGAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

FIGURE 18

MRVRIGLTLLLCAVLLSILASASSDEEGSQDESLDSTTTSDEVKDHTTAGRVVAGQIFLDSEEESEL
ESSIQEEEDSLKSQEGERVTEDISFLESPPNPNKDYEEPKKVRKPALTAIEGTAHGEPCFPFLFLDK
EYDECTS DGDREDGRLWCATTYDYKADEKWGFCETEEAAKRRQMQEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTGFALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATT TAAG CCCATTCTGCAGT GGAATT CATGAACTAGCAAGAGGACACCATCTTCTT
GTATTATA CAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTGGGTGCTAGG
CCTCCTAATCCTCTGGTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTATCACTGGATGTGACTCGGGCTTGGAAACTTGGCAGCCAGAACTTTGAT
AAAAAGGGATT CATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCC GGCGTGCCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGAGGTGCCTTGCAATCGTGGAGGGGCTATACTCCATCCAAA
TATGCAGT GGAAGGTTCAATGACAGCTAAGACGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCCGGTAGAGTG
CATGGACCACGCTTAACAAGTCTCTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTACTCAGTAACCACAAATGTCTCCAGGCTATGA
AATTGGCCGATTCAAGAACACATCTCCTTCAACCCCATTCTTATCTGCTCCAACCTGGACT
CATTAGATCGTGTCTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGTATCCCAGGGT
CCCTGCTCAAGTTCTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGGCCATTCAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGTCCCCAGCATTACAGTAACCTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAA
AAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTAACACACGGACTCGGACACGGAAGGTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAACATTGA
CATTAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTCAAATGTCAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAACTGCAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAACACCAAGTATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAAACCTCAAAAGGACTTTCACAGGGTACCTTAGTGGTT
GCCAATCTGGCATGTCGAACAACGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTGAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTCTTGTCAAGCATTACGGACCTTTCCAAATTCTGAATTCTCATTCACTGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAACCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATCTGGTAGTAGTAACCAAGATAAGCATTCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATATTACGGTCTC
CTACATTTGATCTTTAACCTTACAAGGAGATTTTATTTGGCTGATGGTAAAGCCAAAC
ATTTCTATTGTTTACTATGTTGAGCTACTGCACTTCAAAAGTACTTTCAAACATCA
TGTTGCACTAACACAGATAACTCTTAGTGCATTACTCACAAAGTACTTTCAAACATCA
GATGCTTATTCCAAACCTTTTACCTTCACTAACAGTGGCTCACACCTGTAATCCCAGCACT
ACACATTCTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCAGTCTATTAAAAAATGGAAAAGCAAGAACAGCTTATTGTTAAAGTAACTTTTAGAAGTA
GAAATTATGAAAATTCTGAGTCATTAAATTCTCCTTAAGTGAATGTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGEVKGEAKNSITDSQMDDVEVVTIDIQKYI
PCYQLFSFYNSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVANLGMSEQLGYKTVSGCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVKSSCNYNHLDVVDNLTL
MVEHTDipeaspastPQIIKHKALDDDRWQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGCGGGAGGGCAGAGTCAGCCGAGCCAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACGCCCCGCCGCCACACCCCTCTGCGGTCCCCCGGGCGCCTGCCACCCCTCCCTCCCTCCCC
GCGTCCCCGCCCTGCCCGGCCAGTCAGCTTGCCTGGGTTCGCTGCCCGCGAAACCCCGAGGTACCCAGGCCGCCCTCT
GCTTCCCTGGGCCGCCGCCCTCCACGCCCTCCTCTCCCTGGGCCGCCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTCTACTTTGCCCGCTCTCCGCCGCTCGCTCTCCACCAACTCCAACCTCCCTCTCCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGGCCGCTCCGTCCGGTCCAAA
GGTGGGAACCGTCCGCCCGGCCAGCACCATGGCACGGTCCGGCTTGCACCCCTGGCAGTGCCT
AGCGCCGCGCTGGCTGCCGAGCTCAAGTCGAAAAGTGTCTGGAAAGTGCAGCTTACGTGTCCAAAGGCTTC
AACAGAACGATGCCCTCCACGAGATCAACGGTATCATTGAAGATCTGTCCCCAGGGTCTACGTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGCAATCATTG
CAAGCTGTCTTGCTTACGTTACAAGAAGTGTGAATTCTCAAAGAAACTACTGAAAATGCAGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGCCATTATACATGCAAATTCTGAGCTATTAAAGATCTCTCGTAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGGAAAGAAATGCTAAATGACTTCTGGCTGCCCTGGAGCGGATG
TTCCGCCCTGGTGAACCTCCAGTACCTTACAGATGAGTATCTGGAATGTGTGAGCAAGTACGGAGCAGCTGAAG
CCCTTCGGAGATGCTCCCTGCCAAATTGAAGCTCCAGGTTACTCGTGTGTTGTAGCAGCCGTACTTCGCTCAAGGC
TTAGCGGTTGCCGGAGATGCTGTGAGCAAGGTCTCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGTCTCGTACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTGGCCAACCAAGGGATCTGATTTGAATGGAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGCTTCAACATTGAATCGGTATGGATCCCCTCGATGTGAAGATTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCCCTCCAGCTGGACGAATTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGGACACATCACCCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTACTGATGTCAAGGAGAACTGAAACAGGCCAGAAATTCTGGTCTCCCTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGCCATGAGGATGACTGTTGAATGGAAAGGCCAAAGCAGGTAC
CTGTTGCACTGAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGGTACACCAGCAAACAGAC
ATACTGATCCTCGTCAAATCATGGCTCTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGAACGACGTG
GACTTCTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCAATGAGAAAGCCGACAGTGCTGGTGTCCGCTGGGCA
CAGGCCCTACCTCCACTGTCTTGATCTGCTTCTGGTATGCAGAGAGAGTGGAGATAATTCTCAAACCTCTGAG
AAAAAGTGTCAAAAGTTAAAGGCACCAGTTACCTTCTACCATCCTAGTGACTTGTCTTTAAATGAA
TGGACAAATGTACAGTTTACTATGTGGCCACTGGTTAAGAAGTGTGACTTGTCTCATTGAGTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTCTGCTCCCCAAACCATGTTAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTGTGATTTTACTCTATTATTGTTGTATGTTTTCTCATTGTTGTGGGTT
TTTTTTCCAATGTGATCTGCCCTGTTCTACAAGCAAACCAGGGCCCTTCTGGCACGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTATTATCATGTTATCTTATTAAAGAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKLNDMFVKTYGH
LYMQNSELFKDLFVELKRYYVVGNVNLEEMLNDWARLLEMFRVLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHCRL
VTVKPCYNCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFOQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVFFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCATTATATTCTCAAGCAACT
TACAGCTGCACCGACAGTTGCGGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCAGGCCAAGAATGTGAGTGCAAAGATTGGTCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAACAGCAGTGCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTGTTCAAGATCATTGTTGCTCTCTAGTGTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTACCCAGGCTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTAAACCTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAATATTTT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPDCDFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAACGGGAGCTTCAGCCAG
ATAAAGGGCTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGCGGAACCGGAGC
CAACTTCAAGTCTACCCACCAGGGCTGCAGGGTACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGTGCACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTCTGGAGCATGTGGCAGAACATGGAAAGCTGGCCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGATGGCTGCTGCCTCACAGAGAGACCTGGAAGGATCTTGA
GAACGCCAGTTCTCCAAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGAAAGGCTGTCAAACAATCTTCCAAGCTCCAAGGCACTATTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACTA
GCAGAACATGAGAGAACATTGATGTACCACTACTAGTCCCTCTCTCCCAACCTCTGCCAGGGC
AATCTCTAACCTCAATCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCTGTTGTATCCTCAACTGCAAGTTCTGGACTAGTCTCCAACGTTGCCTC
CCAATGTTGCCCTTCCTCGTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACCTGGTACAAAAGTCATGGTGCCTGCATCCCTGCCAGCCCCCTGAC
CCTCTCTCCCACTACCACCTTCTCGAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTCTCAAATATTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGTCCCAGAAAGTCTCTCTG
CCACTGACGCCCATCAGGGATTGGGCCTTCTTCCCCCTCCTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGAGATAGGAACAGAACAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTATTAAA
GTGGTTTTATGATTCTATACTAATTATAACAAAGATATTAGGCCCTGTTCAATTAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTGAAAGATTGTTCTGTGTAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCQAPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCGAATTCCCTCAACTATACCCACAGTCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTATTAAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTGATAAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGGATCATGTCGG
GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTGTCTTGGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTCTGGTCACATTTCATTGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGTCCTCTGGTGGCTGTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCAGCCAAGTGGATATAAGCCCCTTCGGGCATTGGTACATGTGGTGTACCTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGGTGCAGCAAATGACTATAGCTGGGCAG
TGGTTACTTGTATTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATTCTCGTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGAAAGGGTCATTTAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTCATTGT
CCAGGTACCTGTCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCAAGAATGCATATACTACAACGTCTATTAAATGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTGTCCAAGAACTCAAGTCACTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTCACTGTTGGAGGACTCATGGCTTTACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTGC
CCATAGTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGAATGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTCTAAGAGCCATTACAGAATAGAAGATGAGGACCACTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAAACCTATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGLLFVCGLWWLYDYTNLDSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDPPDHPISSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCYCFCWCLDKYLLHQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCCCTGTTCTCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAAGCCTAAGGGAAAAGAATATTCAATTCTGTGTGGT
GAAAATTTTGAAAAAAAATTGCCTCTCAAAACAAGGGTGTCAATTCTGATATTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTGTTGCTGGTACTGGAGTAC
ATTCAAACAAAGAAACGGCAAAGAAGATAAAAGGCCAAGTTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGGAAAGATCATCGATCCTGAGTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAAGGAGGGAAAATACTGTTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAACCTT
TATCGTCTAGAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCAGCTTACATACATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAAC TGACAGCCGGCACTCTGATGCAGCTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTGCCAAGGCCATCCCCCTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAACTCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTCAGGAGCTGCCCTCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGTCTTGGAGGCCAGTATCCC
TGGGAGATCCAAACTGCAAATGACTTGTGTTTAATTGATGGAGCACCAGCATTGGCAA
CGCGATTCCGAATCCAGAAGCAGCTCTGGCTGATGTTGCCAAGCTCTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTGTCAGTATGGAGACAAACCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCTCCTTGTGACCAAGAACTCTTCCAAAGCCAATGAAACAGAAGCGG
GGCTCCAATGTGGTGGTGGTGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTCATCACCATTGAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGGCCAACTTGCACAAACAAGGCCGTGCAAGAACAAACGGCTTACTC
GCTCACGTGCAGAGCTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCAGGCTGCG
ACACTGACCGCCTGGCCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTCATCGAC
GGCTCCAGCAGTGTGGGACGGCAACTCCGACCGTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGGCATCGGCATCGGGCCGTGCAAGTACACCTACGAACAGCGG
TGGAGTTGGGTTGACAAGTACAGCAGCAAGCCTGACATCCTCAAGCCATCAAGAGGGTGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCAAACAAGAGGAAGTTAATGATCCTCATCACCACGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCGATAGGCCTGGCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACCTCTTGTGGACGA
GTTGACAAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTACAGAGTCAACTCAC
AGCCTCGGAACTGAATTCAAGCAGGAGCAGGAGCACCAGCAAGTGTGCTTTACTACTGACGTGTT
GGACCAACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTGGCAGGGCATGGAGAAC
AAATGTCTGTTATTCTTGCATCATGCTTTTCATATTCCAAAACCTGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCAAAGGCTACATCATGTTGAGGGTGTGGAGATTTCAT
TTTGACAATTGTTCAAAATAATGTTGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTGTGAGATTGTTAAGTGTGTTGAACTCTGTAACCCCTCAGCAAGTTCA
TTTGTCATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAAATAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTYPSALTYSSSKSPAAGETTKAYQRPIP GTTAQPVTLMQLLAVTVAVA
TPPTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQPLVKRVCNDRLACSKTCLNSADIGFVIDGSSVGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKR VGYWSGGTSTGAAINF AQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGTTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTTCATCTGGAATTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCGCTGGCTGCTCTACCGTGGTCTCTCCGACTACTCACCCGAGTGT
AAGAACCTTCGGCTCGCGTCTGAGCTGCTGTGGAGGCCTCGGCTCTGGACTGTCCTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTGCTGCTGTCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTGAACGGATGTACTCTATGAGTATGAGCCGATT
CAGACAAGACTTTCACTTCACACTCGAGAGCATTCAAACGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCCATTAGAGTTACTTGGGTGAAAAAGTCT
TGGTGGGATATGAGGTTCTACATTTCATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTCTTATGGTACATAATCCGACAAGATTTCAGACACATATAATA
ACCTGACCTTGAAAACCATTATGGCATTCAAGGTGGTAAGTGGCTTGCCTTGCCTTGCCTTGCCTT
AAGACAGACACTGATTTCATCAAAACTGGCAATTAGTGAAGTATCTTAAACCTAAACCAACTCAGA
GAAGTTTCACAGGTTACCTCTAATTGATAATTTCCTATAGAGGTTACCAAAAAACCCATT
CTTACCAAGGAGTATCCTTCAAGGTGGCTCCACTGCAGTGGTTGGTTATATAATGTCCAGAGAT
TTGGTCCAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAAGACACAAATCTTCTTCTATAGAATCC
ATTTGGATGTCTGTCACTGAGACGTGTGATTGCAGCCATGGCTTCTCCAAGGAGATCATCACTTT
TGGCAGGTCATGCTAAGGAACACCACATGCCATTACTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTAAATAAGTAGGTACTGTGGAAAATTCAATGGGAGGTCAAGTGTGCTGGCT
ACACTGAACGTAAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTGTGATTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAAATAAGGAATTGGAGGTTTGCTAAAGAAATTAAATAGG
ACCAAACAATTGGACATGTCATTCTGTAGACTAGAATTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAATGTAGAGTTATTATGAAACAATGTAGTCATTGAAAGGTTGTGTA
TATCTTATGTGGATTACCAATTAAATAATATGTAGTTCTGTGCAAAAACCTCTTCACTGAAGTTATA
CTGAACAAAATTTCACCTGTTGGTCATTATAAGTACTTCAGATGTTGCAGTATTCACTGAGTTATT
ATTATTAAAATTACTCAACTTGTGTTAAATGTTGACGATTCAACAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTAACAGTCAAGTATTGATAACATCACTCCA
TTAATGTAAAGTCATAGGTCAATTGCATATCAGTAATCTCTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNCQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLISSLPREYTVINEACPGAEWNIMCRECCYEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRCVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTAQQNGEWSGKQPIKACREP KISDLVRRVLPMQSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLOYECISPFYRRLGSSRTCLRTGKWSGRAPSCIPICGKNIENITAP
KTQGLRWPWQAAIYRRTSGVHD GSLHKGAWFLVC SGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDRDEKTIQSLQISAIILHPNYDPILL DADIAILKLLDKARI STRVQPI CLAASR
DLSTS FQESHITVAGWNVLADVRSPGFKN DTLRSGVVS VVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPS DICTAETGGIAAVSF PGRASPEPRWLMGLVWSWSYDKTCSHRLSTAFTKVL PKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCATACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCGTGATTTATTACGTGGCTTAATCTGAGGTTCTCAGTC
AAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGAGAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCTGGCTCAGTCCTGCTAACTACATTGACAATGTGGCAACCTGCACCTCCTGTATTCAA
GAGACTCTGTA
AAGGTGCCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCCTCCTTAATGACAGACGGCTGGCTAGACA
ACCCCTGCCTACGTGTCCTGGCAGAGGACGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGA
GGGCACGGCCCTTGAGAGATCCACTATTAGAACGAGTCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAA
ACCAACTGCCCTGAAG
TCTTCCAAGGTTGTACCACCTGATTCCAGATGGTAAATTACAGCATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGAGGTAGCGAAACCCCCTGCTCATATCATTATCCAACACATTTATCGTGTGG
TGATGCCAGAGACGGCCGGCTACTGCCAGGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCC
ACAACACTACGCTGTGCGTCTCCTGCCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCC
GCAGGAACAATGGACAGGCCCGGATGCCCTACAGACCCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAAACTGGTGCAGCAAGGTGGATGAGCCTGGGTTTCATCTTCAATGTGCTGGATGGC
GTGTGGCATATCGACATGGTCAGTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCAGAAAGTGCAGCTCATCTGATTCAAGGCCAGTGAAAGACGTGTTCACCTCGTGTCCGCCAGGTTGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGGGAGCAACA
CTCCCAAGGCCCTCCATCCTACAAATTACTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCGGTGAATCTCTG
GCATGACCGTCGCAGGGGGAGCATCACATAGAGAATGGGATTGCTATCTATGTCATCAGTGTGAGGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTGTGAAATGTGGATGGGTCGAACGTGAGAGGTGAGCC
GGAGTGGCAGTGGCATTATTGAAAGAACATCATCCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAAACATGGCCCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACACGGTGTATAACTGTAAGATATTGTATTACGAAGAACACAGCTGGAGTC
TGGGCTTCTGCTGATTGTAGGAGTTATGAGAATAACATGGAAACAAACCTTTTCAAAATCCATTGTTGAAGGAA
CACCAGCATACATGAGAAGAATTAGATGTGGTGTGATATTCTCTTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGTGGCAAGACTGCTGAAAGAACTTAAAGGAGAATTACTCTAAACTATTGTTCTTGGCTGGCACTT
TTTTATAGAATCAATGATGGTCAGAGGAAACAGAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTATC
TTGTCAGTTTATATTAAAGAAAGAACATTGTAAGGAAAGTGTGAGGAAACTGATCATCTAAATGAAAGCCAGTT
ACACCTCAGAAAATGATCCAAAAAAATTAAACACTAGTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATATTCTATTCAATAAAAGCCCTAAACAACTAAAATGATTGATTGTATAACCCACTGAATT
CAAGCTGATTTAAATTAAATTGGTATATGCTGAAAGTCTGCCAAGGGTACATTATGCCATTTTAATTACAGCT
AAAATTTTAAATTGCTGAGAACGTTGCTTCAAAACAAGAATAAATTTTCAAGAAGTTAAA

FIGURE 40

MKALLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGSRNRTRARPFERSTIRSRFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKVDPEGVFI FNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVVSQRQSPDI FQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGVISRDGR
IKTGDILLNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAACATGGGCTAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTGTGCCTGACAGTGGTTGGGTGGC
CACCACTAAGTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCTCATTTGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACCT
GACAACTGTCCTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAAG
CTTACAGAGGGTCGCCATCCTCGTTCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCCAAACTCTTGAATGTGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTGG
GGGTGTTACTGCCCTAACGCAGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCGGCCCTG
CCTGAAGTGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGTTGAGTAGTTGTTCTATAAAT
TAGTATCTGTGGAACCACAATCCTTATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTATGTTGGAAGAACTGATTCTTGGTCAATAATTGGCTAGAGACTCAA
ATAGTAGCCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTCT
TAGCAGAGCTCCTGGTATGTAGAGTATAAACAGTTGTAACAAGACAGCTTCTTAGTCATTGAT
CATGAGGGTAAATATTGTAATATGGACTTGAAGGACTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTTAAATTGAGTAATATATTATGGGAT
AAAAGGCCACAGGAATAAGACTGCTGAATGTCTGAGAGAACCAAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCAATTACCTGTACAATCATCTGTGAAGTGGTGGTCAGGT
GAGAAGGCGTCCACAAAGAGGGAGAAAGGCGACGAATCAGGACACAGTGAACTTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCTCCAGTGAATGCCACCAGAGAATACATTCTATTAGT
TTTAAAGAGTTTGTAATGATTGACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACTAATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQOLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVITALSREQFFKVNNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGTGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACTTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGGAAATAAAATGGCTGGTTTTGTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCGTGAAAGAGCTGGTCGGT
TCCGGTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAAACCCCTTTGTCACCACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGTATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTTTGACTGGGGCTATTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTGTCGGG
AAACTCCTAACATATGCCCTCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAAATCCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGGAAATACGAAAAAGAT
GGAAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTCTCAAAACCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTGACAGCAAGAAAAGAGCGTG
CTGCAACAAACAGAACTGGAATGTTCTTCATCATTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAGGTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATTGAAAAACATCAGTGACATTCCAGAATCCTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGCACCTCCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTTCAGGTCTATTGCTTGGATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTAAGCGAAGAAGTCAAATTG
TGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLGVVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCGCCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCTTCCAAGCCAATGACC
CATTGCCAATAAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCTGGCCATTGCTGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACGAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCTGAAGCCTAACACTGGCCCCAGCACC
TCCTCCCTGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGACAAATATTGGGAGGCCCTGGACATGCCCTGGAGACGCCCTGA
GCGAAGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCGC
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCCATGCTCTGGAAACACTGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGGCTCCTGGCAGGG
GTGCCCTGGCACAGTGGTGGAAACTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGCAATGGTACAACAAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGCAGTGGCAGCAGCAGCAGTGGCGGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGCAGTGAGT
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCGGGAGGAAATGG
CATAAACCCGGGTGTGAAAAGCCAGGAATGAAGCCCGGGAGCGGGGAATCTGGGATTAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGCAATGCCCTCTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCAGCTGGGAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTGAGACGTCTCCTGGATGTTAACCTTGACACTTCTGGAA
GAATTAAATCCAAGCTGGTTCATCAACTGGATGCCATAAACAGGACCAAGAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGGCCACACTCCCTCTAA
AACACCACCCCTCATCACTAATCTCAGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGP
GHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGGSSNSGGSGSQGSSGSGSNGDNNNGSSGGS
SSGSSSGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSSGNHHGSGGGNGHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCTCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACAGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACACTGGTTGGGTACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCCAT
CATCCCCTTCATCGTTATGCCACCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGAACGCAAGCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCCATGACGGCGCGCTTCCACAGGGCTGCCGCTGGTCATGACTCACAGAC
GCTGTCATCCGGAGCGCGTGCACCCCTCCCCACTCAGGGTATTGATGATTAAAGACAA
AGCCAAGTCCAAGACTTGGATTCATTGATGTGCTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCAGGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGACGACCTGGCCC
AGCTGCCCTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTACATCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATCACAACCCAACTGTGTCGGCGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAACAGCAAGGGAGGTACCTCTGGCTTTATTCTTCTCCGCA
GGGCCAGGAACACTGCATGGGCAGCGTGCCTGCCATGGCGAGATGAAAGTGGCCTGGCGTTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCAACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCAGATTTTCG
CCATCCACCTGTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWGLRPVAMSPWLLLLVVGSWLLARILAWTYAFYNNCRLQCFPQPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGLLGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHGRRFHRAC
RLVHDFTDAVIRERRTLPTQGIDDFFKDKAKSKTLDFIDVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPFVDPFRDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGCTGTTGTTGCACTTACCCGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTGCACTTATCTTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTTGC
ATAATTCATGCCAGTTATGAAGCTTGAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACCTCTGTCTTACAGACATGTGCCTTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGTTCTGTATCTTGTGGAGTGGAAATCTTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCCAGTGAATTCCCATGAATACAAACCTATTCAACAGCAACAGCAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGAGTGGAGCACCCAGCAGGCCCAACATGCTCTGTGCCTG
TACGTGCCGGTCACTGGGAAGCCCAGACCGAGTTCCAGTACTTGAAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCACCTACCGCCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTGAAGAATTGTCCATTATCTCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTGGACAAAAGAATGATGGACGATGACGCCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAGATACTGAACAGCAGGAGAAAAATTCTCAAGAGCATGGATAAAACGGCACGATGACCATGACTGG
AACGAGTGGAGAGACTACCACCTCCACCCGTGGAAACATCCCCGAGATCATCCTCTACTGGAAGCATCCACG
ATCTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGGAGGAGGGCAGACGGGATGTGGTGGAGA
CACCTGGTGGCAGGGAGGTGGGAGGGCGTATCCAGAACCTGCACGCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCATGCCCTCCGCAGCAACACATGGGCATCGTGGCTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGCAATGGCATCAACGTCTCAAAATTGCCCGAATCAGCCATCAAATTATGGCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGCAGGAGACTCTGAGGATTACAGAGAGGCTGTGGCAGGGCTTGGCAGGGCC
ATCGCCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCCTGCCAGAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCCAGGAGGATCTGGCAGAGAGGGGTGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTG
GGCATCATCCCTATGCCGCATCGACCTTGCAGTCTACGAGACGCTCAAGAATGCCGGCTGCAGCACTATGCAGT
AACAGCGGGACCCGGCTGTTGTCTGGCTGTGGCACCATGCCAGTACCTGTGGCAGCTGCCAGCTAC
CCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGC
TTCAACATATCTGCGGACCGAGGGGCGCTTGGGCTGTACAGGGGCTGGCCCAACTTCATGAAGGTATCCA
GCTGTGAGCATCAGTACGTGGCTACGAGAACCTGAAGATCACCTGGCGTGCAGTCGGTGTGACGGGGAGGGC
CGCCGGCAGTGGACTCGCTGATCCTGGCGCAGCCTGGGCTGTGAGGCTACAGCCATCTCATTCTGTGAATGTGCC
AAAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGTGGGAGAGCTGGCAGGCCAGGGCTT
GTCTGCTGACCCAGCAGACCCCTGTGGTCTACGAGAACCCAGGCTTACAGGCTTACAGGCTTACAGGCT
CTCCGGCTCACATGTGAAGGACAGGACATTTCAGCTGAGCTGGCCATGGCTGGAGGCGCT
TAGTTCTCCATTTCACCCCTGCAGCAGCTGTGGCCACGGCCCTGCCCTGTGCTGCCGTGCATCTCCCTGTGC
CCTCTGCTGCCGTGCTGAGGTAAGGTGGGAGGCTACAGCCACATCCCACCCCTGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCTCCAGCCTGACTTCCAAACCTACAGCATTGACGCAACTTGGC
TGTGAAGGAAGAGGAAGGATCTGGCTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGGAGGG
CTTGGGAGTGCAGGGGCTCGGCTGCCCTGGCTGCACAGAGGCAAGTGTGGGCTCATGGTGTCTGAGCT
GGCCTGGACCTGTCAGGATGGCCACCTCAGAACAAACTCACTGTCCCACGTGGCATGAGGGCAGTGGAGCA
CCATGTTGAGGGCGAAGGGCAGAGCTTGTGTTCTGGGAGGGAGGGAAAGGAAAGGTGTTGGAGGCCTTAATTATGG
ACTGTTGGAAAAGGGTTTGTCCAGAAGGACAAGCCGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGCTCTGTCCAACCCAGCAGGGCGCAGC
GGGACACCCCCACATTCACTTGTCACTGCTGGAACCTATTATTTGTATTATTTGAACAGAGTTATGTCT
AACTATTTTATAGATTGTTAATTAAATAGCTTGTCAAGTTCAAGTTCAATTATTTTATTATTTATGTTCATGGT
GATTGTACCTTCCAAGCCGCCAGTGGATGGGAGGAGGAGAAGGGGGCTTGGCCGTGCAGTCACATCT
GTCCAGAGAAATTCTTTGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTGGCTCCTTCTGGCAG
GTTGGGAGGGCTTGGCCCTAGGATTCTAGGTTGACTGGGGCTGGAGAGAGAGAGGGAGGAACCTCAAT
AACCTGAAAGGTGAATCCAGTTATTCTGCGCTGCGAGGGTTCTTATTTCACCTTTCTGAATGTCAAGGCAG
TGAGGTGCCTCACTGTGAATTGTGGTGGCGGGGGCTGGAGGAGGGTGGGGGCTGGCTCCGTCCCTCCAGC
CTTCTGCTGCCCTGCTTAACAATGCCGCCACTGGCACCTCACGGTGTGACTTCCATTCCACAGAATGACCTGA
TGAGGAAATCTCAATAGGATGCAAAGATCAATGCAAAATTGTTATATATGAACATATAACTGGAGTGTCAAAAAG
CAAATTAAAGAAATTGGACGTTAGAAGTTGTCAATTAAAGCAGCCTCTAATAAAAGTTGTTCAAAGCTGAAAAAA
AAA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTTCCCCAGCCATGGC
TTCCCTGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTGAAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTTTGGCCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCGG
AACTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCATGGTCCCC
CAGCCCACAGTGGCTGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGTGT
CTCTTCTTCTTGCCATCAGCTGGCACTTCTGCCCTCAGCCCTACCTGATGCTAAATAAAT
GTGCCCTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAGATATGACCTAGTTATATTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTGATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCGCCTGT
CACCTGGGAGTGAGAGGACAGGATAGTGCATGTTCTGTCTGAATTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCTGAAAGTCATCCAAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCCCTAACAGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTATGATGCTTCCAAAGGTGCCTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAAACTGAGCACCTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLGVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQOPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIIRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAACTACCCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCTCCGGTAGCTACTACCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTTATGGGGCTGCCCTGGAGGGCCTATGGACCACCAGCTGG
TGGAGGGCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCGCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTTGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTCTAGAGTGAGAGAAGTATGTGGACATCTCTCTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTGCTTGATGCAACACTGTTAAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTAGTGTCTCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAAGGCATCCTGTCACAGGCCATTCTCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTCTAGCTGCCTGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGCCT
CTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTGCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCACTTTTTC
ATTGGGCCAAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSPGTGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTTTGCCCGAGCATCGGTGTTGGAGGAGAAAGTTCCAAAACTCGGGACC
AACTGCCTCAGCTCGGACAACCTCCTCCACTGCCCTCTAACACTGAACATCCGAGCCC
TCTGGACCCTAGGTCTAATGACTGGCAAGGGTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATT CCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGG
AGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGCTGCCGCCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGAGTCTCTCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAACACTGGGAGCCGGGGAAAAATCTTCCAAACG
CCCTCCCTGGTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCCCTGAATCCCAGTG
TGT CCTGGGAGGTGGAGGCCCTGGACTGGTGGGAACGAGGCCATGCCACACCCCTGAGGGA
ATCTGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGGTATCCAGGAGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCAATCTATACCCAGGTATCAATAACCCATTCCCTGGAGTTCTCCGC
CCTCCTGGCTCTTGGAACATCCCAGTGCTCCCTAACCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCCTGCTCCGCCCTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCAGTACCTCCCCACTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGGINNQPPGTWGNINRYPGGS
WGNINRYPGGSGNINRYPGGSGNIHLYPGINNPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCCCTCTTGCACCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCCTTGGCCC
GGGCTTTGGCCGGGATGCAGGAGGCAGGCCCGACCCGTCTTCAGCAGGCCACCCCTC
CTGAGTGGCAATAAATAAATTGGTATGCTG

FIGURE 66

MGSGPLVLLTLLGSSHGTGPGMTLQLKESFLTNSSYESSFLELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGAGGTGCCCTGGCGGAGATCGGTAGGAGGGCGAGCGCGAGAAGCCCC
TTCCTCGCGCTGCCAACCGCCACCCAGCCCATGGCGAACCCGGGCTGGGCTGCTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGCCGAGCCTGGGGCAAATACAGACCACCTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGCTCAGGAGACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCTCCTGCATCTGTCTCCCTCATTGCTGTGACCTGGGAAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGAAAGAAGGTACTTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPILLARWGRAWGQIQTTSANENSTVLPSSSSDGTLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GGCAGGAATAACTAGAGAGGAACAATGGGGTTATTCAAGAGGTTTGTCTAGTTCTGTGCCTGCTGCACCAAG
TCAAATACTTCCTCATTAAGCTGAATAATAATGGCTTGAAGATATTGTCATTGTTAGATCCTAGTGTGCCAGAA
GATGAAAAAATAATTGAACAAATAGAGGATATGGTACTACAGCTTCTACGTACCTGTTGAAGCCACAGAAAAAGA
TTTTTTCAAAAATGTATCTATATTAAATTCTGAGAATTGAGGAAATCCTCAGTACAAAGGCCAAACATGAA
AACCATAAACATGCTGATGTTAGTTGACCCACCTACACTCCCAGGTAGAGATGAACCATAACCCAAGCAGTTACA
GAATGTGGAGAGAAAGGCAGAACATTCACCTCACCCCTGACCTTCTACTGGAAAAAAACAAATGAATATGGACCA
CCAGGCAAACGTGTTGTCCATGAGTGGGCTCACCTCCGGTGGGAGTGTGATGAGTACAATGAAGATCAGCCTTC
TACCGTGCTAAGTCAAAAAAATCGAACAGGTTCCGAGGTATCTCTGGTAGAAATAGAGTTATAAGTGT
CAAGGAGGCAGCTGTCTAGTAGAGCATGCAGAATTGATTCTACAACAAAATGTATGGAAAAGATTGTCATTCTT
CCTGATAAAAGTACAAACAGAAAAGCATCCATAATGTTATGCAAAGTATTGATTCTGTGTTGAATTGTAACGAA
AAAACCCATAATCAAGAACGCTCAAGCCTACAAAACATAAAAGTCATAATTAGAAGTACATGGAGGTATTGCAAT
TCTGAGGATTAAAAACACCATAACCCATGGTACACCCACCTCCACCTGTCTCTCATGGTAGAAGATCAGTC
AGAATTGTCGCTTAGTTGATAAGTCTGGAAGCATGGGGGTAAGGACGCCAAATCGAATGAATCAAGCAGCA
AAACATTCTGCTGCAGACTGTTGAAATGGATCCTGGGTGGGATGGTCACTTGATAGTACTGCCACTATTGTA
AATAAGCTAATCAAATAAAAGCAGTGTGAAAGAACACACTCATGGCAGGATACCTACATCCTCTGGAGGA
ACTTCCATCTGCTCTGGAATTAAATATGCATTCAAGGTGATTGGAGAGCTACATTCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTGTATTGATGAAGTGAACAAAGTGGGCCATTGTCAT
TTTATTGCTTGGAAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTGTT
TCAGATGAAGCTCAGAACATGGCTCATTGATGCTTTGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTTCAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAATTGATAGTACA
GTGGGAAAGGACACGTTCTTCATCACATGGAACAGTCTGCCTCCCAGTATTCTCTGGATCCAGTGGAAACA
ATAATGGAAAATTTCACAGTGGATGCAACTTCCAAAATGGCTATCTCAGTATTCCAGGAACGTGCAAAGGTGGCACT
TGGCATAACATCTCAAGCAAAGCGAACCCAGAAACATTAACATTACAGTAACCTCTCGAGCAGCAAATTCTCT
GTGCCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAACAGTTCCCCAGCCAAATGATTGTTACGCAGAA
ATTCTACAAGGATATGTACCTGTTCTGGAGCCAATGTGACTGCTTCATTGAATCAGAAATGGACATAACAGAAGTT
TTGGAACTTTGGATAATGGTGCAGGCCTGATTCTTCAGAATGATGGAGTCTACTCCAGGTATTACAGCATAT
ACAGAAAATGGCAGATATAGCTTAAAGTCTGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAATTACGGCCTCA
CTGAATAGAGCCCGTACATACCAGGCTGGTAGTGAACGGGAAATTGAAGCAAACCGCCAAGACCTGAAATTGAT
GAGGATACTCAGACCACCTGGAGGATTTCAGCCGAACAGCATCCGGAGGTGCATTGTTGATCACAAGTCCAAAGC
CTTCCCTGCCTGACCAATACCCACCAAGTCAAACATCACAGACCTTGATGCCACAGTTCATGAGGATAAGGATTATTCTT
ACATGGACAGCACCAGGAGATAATTGATGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAGTATTCTT
GATCTAAGAGACAGTTGATGCTCTCAAGTAAATACTACTGATCTGTCACCAAGGAGGCAACTCCAGGAA
AGCTTGATTTAAACAGAAAATCTCAGAACCCACATATTGCTTAAAGTAAAGTATAGATAAA
AGCAATTGACATCAAAGTATCCAAACATTGACAAAGTAACCTTGTTATCCCTCAAGCAAATCTGATGACATTGAT
CCTACACCTACTCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTCTACGCTGGTATTG
TCTGTGATTGGCTGTTGTAATTGTTAAGTACCAACCTTGAACCTTAACGAAGAAAAAAATCTC
AAAGTAGACCTAGAAGAGAGTTAAAAACAAATGTAAGTAAAGGATATTCTGAATCTTAAACATCCCAT
GTGTGATCATAAAACTCATAAAAATAATTGATGTCGAGGAAAGGATACTTGATTAATAAAACACTCATGGATA
TGTAAAAACTGTCAGATTAAATTAAGTATTGTTATTGTTATTGTAAGAAATAGTGTGAAACAAAG
ATCCTTTTCAACTGATAACCTGGTTGATATTGATGCAACAGTTCTGAAATGATATTCAAATTGCA
GAAATTAAAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGGAGCAAATAACACATTGGAAAAAA
AAA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTASTYLE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENKHADVIVAPPTLPGRDEPYTKQFTECGEKG
EYIHFTPDLGGKKQNEYGPPGKLFVHEWAHLRGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVT
PPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFL
LQTVENG
SWVG
MVHF
DSTATIV
NKLIQIKSS
DERNT
IMAGL
PTYPLGG
TSICSG
IK
YAFQV
IGELHS
QLDGSE
VLLTD
GEDNT
ASSCIDE
VKQSG
AIVHF
IALGRA
ADEAVI
EMSKITGG
SHFYV
SDEAQ
NNGLID
AFGALT
SGNTD
LSQKSL
QLES
KGTL
NSNA
WMND
TVIID
STVG
KDTFFL
ITW
NSLPP
SISL
WDPSG
TIMEN
FTVDAT
SKMAYL
SIPGT
AKVGT
WAYNL
QAKAN
PETLT
ITV
TSR
AAN
SSV
PPITV
NAKM
NVD
NSFP
SPMIV
YAEIL
QGYV
PVL
GAN
VAF
IESQ
NQH
TEV
LELL
NGA
GAD
SFKND
GVYS
RYFTAY
TEN
GRYSL
VRA
HGG
ANTAR
LKL
RPPL
NRA
AYIP
GWVN
GEIE
ANPP
RPE
IDE
DTQ
TLED
FSRT
ASGG
AFV
VSQ
VPSL
PLD
QYPP
SQT
DLD
ATV
HED
KII
LT
WT
TAP
GDN
FDV
GKV
QRY
IIR
ISAS
IL
DLR
DSF
DD
DAL
QV
NTT
DL
SP
KEA
NSK
ESFA
FKP
EN
ISE
EN
ATH
FIA
I
KS
ID
KSN
LTS
KVS
NIA
QV
T
L
FIP
QAN
P
DD
ID
D
PT
PT
PT
PT
P
DK
SH
NS
GV
N
I
ST
L
V
L
S
V
I
G
S
V
I
I
V
N
F
I
L
S
T
T
I

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCTCCTATTCAAGGAAAGACGCCAAGGTAATTGACCCA
GAGGAGCAATGATGTAGCCACCTCTAACCTTCCCTCTGAACCCCAAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCCTCGGCTTAACCTGTGGTGGAGGAGAGAACCTTGTGGGGCTGCCTCTCTAGCA
GTGCTCAGAAGTGAECTGCCTGAGGGTGGACCAGAAAGAAAGGAAAGGTCCTCTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGAATGAAGGTAAAACCTGGAGATTCACTCAGTCATTGCTCTGCCTGCAAGATCATCCTTAAA
AGTAGAGAAGCTGCTCTGTGGTGGTTAACCTCAAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGCCCAAACGCATGCTCCTGTGGTCTAGGCCAGGGAAAGCCCTCCGTGGGGCCCGCTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTGCCTGGGGCTGCTTGCCTGGATTTCCGGTGGT
GTTTGCTGGTGCCTCTGCTGTGCTATCTGCTCTGATGTTGGCCTGCACCCCAAAGGTGACGAGGAGCAG
CTGGCACTGCCAAGGCCAACAGCCCCACGGGGAAAGGAGGGTACCAAGGCCCTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGCGGAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGCAGCTCAGG
AATGGCAGTACCAAGCCAGCAGTGCCTGGCTGGGCTGGACAGGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTCCTGCACTCGCAGGTGGACAAGGCAGGGTGAATGCTGGCCTGAAGCTGGCCACAGAGTATGCAGCAGTG
CCTTCGATAGCTTACTCTACAGAAGGTGTACCAAGGAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAACGCTGT
AGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTGGAGACCCCTGAACAATCCTGCAGAGAACAGC
CCCAATCACCGTCTTACACGGCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAGGGACATTGTAT
GAGCTACCTTCAAAGGGACCACAAACACGAATTCAAACGGCTCATCTTATTCGACCATTAGCCCATCATGAAA
GTGAAAATGAAAGCTAACATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAGAATTCAAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCATCTCACTGTTGTTACTTGGG
AAAGAAGAAATAATGAAGTCAAAGGAACTTGAAAACACTTCAAAGCTGCCACTTCAGGAACATTACCTTCATC
CAGCTGAATGGAGAATTCTCGGGAAAGGGACTTGATGTTGGAGCCCTCTGGAGGGAAAGCAACGTCCCTC
TTTTCTGTGATGTTGGACATCTACTTCACATCTGAATTCTCAATACGTGAGGCTGAATACACAGCCAGGGAAAG
GTATTATCCAGTTCTTCAGTCAGTACAATCTGGCATAATACGGCCACCATGATGCAGTCCCTCCCTGGAA
CAGCAGCTGGTCTAAAGAAGGAAACTGGATTGGAGAGACTTGGATTGGATGACGTGTCAGTATCGGTAGAC
TTCATCAATATAAGGTGGGTTGATCTGGACATCAAAGGCTGGGCGGAGGAGTGCACCTTATCGAAGTATCTC
CACAGAACCTCATAGTGGTACGGACGCCTGTGCAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCAAAGGCATGAACGAGGACATCCCACGGCAGCTGGCATGCTG
GTGTTAGGCACGAGATAGAGGCTCACCTCGCAAACAGAACAGAACAGAACAGAACAGAACATTGAACTCCCAGA
GAAGGATTGTGGGAGACACTTTCTTCCTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCATAAA
GGACGACAAAAGAATTGGACTGATGGGTAGAGATGAGAAAGCCTCCGATTCTCTGTTGGCTTTTACAACAGA
AATCAAATCTCGCTTGCTGCAAAGTAACCCAGTTGACCCCTGTGAAGTGTCTGACAAGGAGAACATGCTTG
AGATTATAAGCTAATGGTGGAGGTTGATGGTGTAACTACACTGAGACCTGTTGTTGTGCTCATTGA
AATATTGATTTAAGAGCAGTTGTAAGGAAATTCACTGATGAAAGGCAAGCATATTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAACGGCAGGAGAGGAGATAGGCTTATTGATACT
AGTGAGTACATTAAGTAAATGGACGAGAAAAGAACATAAATATCGTGTCTATTTCCCAAGAT
TAACCAAAAATACTGCTTATCTTTGGTGTCTTTAACTGTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
TTTTTCCCTGAGTTAGTCTGCTTATTAACTTACCACTTGCAAGCCTACAAGAGAGACAAAGTGGCTAC
ATTTTATATTCTTAAGAAGATACTTGAGATGCATTGAGAACTTCAGTTCAAAGCCTACAATTGATGCCATAT
CCAAGGACATGCAAATGCTGATTCTGCAGGCACTGAATGTCAGGCAATTGAGACATAGGGAAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTCTGAAGAGTATTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTTCTGCTTACAGAAAAGGAAACTCATTCACTGGTATCGTGTACCTAAAGTCAGAAACCACATT
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAAAGTATACCGTGTGAACCAAACAATCTTTTC
AAAACAGGTGCTCCTCTGGCTTCTGGCTTCAAAGAACAGAAAATGGAGAAAATATATATATATATATATTGT
GAAAGATCAATCCATCTGCAGAATCTAGTGGGATGGAAGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAG
TAAGTGAATTATTAAAGCAGTTACTCAATCAGAACATAGAGTGGTTCTTCATTGAGAACATTAGCCAGCAC
CAAACATTAAAGAACAGAACATAGAGTGGTTCTTCATTGAGAACATTAGCCAGCAC
ATGCATGAGCTAATTATCTCTTGAGTCCTGCTGTTGCTCACAGTAAACTCATTGTTAAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAATGCATTGATTGATTGACTGGTAGTTATGAAATTAAATTAAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTGTTGAGATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNELNMAN
TLINVIVPLAKRVDKFRQFMQNREMCEIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFPVLFSQY
NPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTQCYRSDFINIGGFDLIDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGAGGGCAAAGAGGCAGCAAGAGAGTTGCCTGGGATCCA
GAAACCCATGATACCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCTGTCCTAGTCCTAGTCCTCAAATTCCCAGTCCCCTGCACCCCTC
CTGGGACACTATGTTCTCCGCCCTCGCTGGAGGTATTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGCCAGCCTTACCCGTAGTGT
GGAAACAATGCCAGTCGCCATCGATATTGACAGACAGACTGTGACATTGACCCGTATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTTACCCGTATCTGGGTGACTTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACGGTCAGAAAGGATCCCCAGGGGGTCAGAACACCAAGATCAACAG
TGAAGCCACATTGCAGAGCTCCACATTGTACATTGACTCTGATTCTATGACAGCTGAGTG
AGGCTGCTGAGAGGCCAGGGCTGGCTGTCCCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAGAGTGCTCTGGACAGTTTTATAGAAGGTCCCAG
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCCACAGAACAGGAGCCCTCAA
GCTTCTGGTACAGAACTACCGAGCCCTTCAGCCTCTCAATCAGCGATGGCTTGTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTGGTTGGC
TGTCTGCCTCTCGTGTGTTATTGCTAGAAAGATTGGAAGAACAGGCTGGAAAA
CCGAAAGAGTGTGGTCTCACCTCAGCACAGCCACGACTGAGGCATAAATTCTCTCAGATAAC
CATGGATGTGGATGACTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTGTTAGTTGCAGGGGAAGTTGGG
ATATACCCCAAAGTCCTTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATTTGATCAATATATTGGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGACGGGAGTTCCCTGTGTC
TCTGGTGGTTGCCTAACCTGCAAACATCACCTTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGGTCTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAACAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTGCCATATCTATTACCGTGTCTTTT
TCTGTGATGGCTATTCCATCTACGATATCCACGTTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTATGAAATGAATTGACAAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGATTA
ACTTTATCACCTCAATATCTGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGTATGCTCGCATTGATGAAATTGGTACTCTGAAG
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTGGAGTCGCAGGCAGCGTTGGCAG
TCTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCTGGTCAGTGGGA
TCCCCAAACTGGCAGGCTGTATTCCCTCGCTGTCAGCTCGACCAGGATTGAGGGCTGCG
AGCCTCTGAGGGGGATGGCTCGGAGAGGGCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAATGAAACCTATCTCATGCAATTGAGGAATGGGGTTATA
TGTGCAGATGAAACTGATGCCAACACTCCTTGCCTTGTGCAAACAAGTGAG
TCACCCCTTGATCCCAGCCATAAGTACCTGGATGAAAGAAGTTTCCAGTTGTCACTG
CTGTGAGAATTACTTATTCTTCTATTCTCATAGCACGTGTGATTGGTTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGCTGGCCGGTTGTCTATGCAGAGAA
AGCAGTCAATAATGTTGCCAGACTGGGTGCAGAATTATTAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSRTWSQCVTNHTLVLWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIGNEFDKRFFVPAEK
IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVKHTMLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOEEVSTQGTLLESQA
ALAVLGPQLQYSYTPQLQDLDPLAQEHTDSEEGPEEPTTLVDWDPQTGRLCIPSLSFDQDS
EGCEPSEGDSLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMen

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTCCCCAGGTCTGGCATCCTGCACTTGCTGCCCTTGACAC
CTGGGAAGATGGCCGGCCGTGGACCTTCACCCCTCTGTGGTTGCTGGCAGCCACCTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGTGCTCAGTGCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAACCCATCCGCATGGACACCAGTGCAGTGGC
CCCACCCGCCTGGCCTCAGTGACTGTGCCACCAGCCATGGAGCCTGCGCATCCAACTGCTGTA
TAAGCTCTCCTCCTGGTGAACGCCCTAGCTAACAGGTCAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTAAAAACCAAGCTGTGCTCCGTGATCGAGGCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCATTCAGCTCTACCTGGGGCCAAGTTGGACTCACAGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTAAAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCATGGCTGTGGACTCTGTGCTTCTGAGAGTGCCATGGCTGAAGTCAAGCATGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTAAGATCTAACTCAGGAC
ACTCCCGAGTTTATAGACCAAGGCCATGCCAAGGTGGCCAACTGATCGTGTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTACCCCTGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAAATAACATCAGCTCTGATGGATCCAGCTG
ATGAACTCTGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCAACAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTCGAGGCAGTGAGTCCTCACTGACCAAGGATGCCCTGTGCTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTGGATGGCAGCCATCAGGAAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCAATCCCTCTGCAATCAAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIWLVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPMTLDNIPFSLIVSDVVKAAVAAVLSPEEFMVLLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSPPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTCTGGGGCTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGGATT
ATT CCTGTTGCCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTTTACTTGGCATTATTCTCCCTGTTCCCTGATAG
CTGGAATCATCCTCTGCTTTCCCTGCTCATCCCAGAGAAATCGCTCAAACACTACGATGCC
CAAGCCCCAACCTCTGCCACAAGGAGCTCTCAAGGCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCCCTACAGCCTGACAGGGTATGTGTGAAGAACCAAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTAAAAAACAGTGGACAGCACCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTCC
TCACCTGCTGCTCCCTGCCCTAACGCCCCAACCTCAACTTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCTCTGGTTACCTGGACTCCATCCCCAACCCACTAATCACA
TCCCACGTACTGACCCCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCCACCTTGTATGACTCCACAGTGTCCA
GACTAATTGTCATGAACTGAAATAAAACCACCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGAGCAGCCTGGGACATTAAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACGCGTCCGCGCCTCTCCCTCTGCTGGACCTTCCTCGTCTCTCCATCTCTCCCTCCTTC
CCCGCGTTCTCTTCCACCTTCTCTTCCACCTTAGACCTCCCTGCCCTCCTTC
GCCCACCGCTGCTTCCTGGCCCTCTCCGACCCGCTCTAGCAGCAGACCTCCCTGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTCGTCTCCCTCCGACTCCGCTCCGG
ACCAGCGGCCTGACCCGGAAAGGAGTGTTCCGAGTGAGGTCTCTGTGG
CTCGCGCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTGGAGGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGGCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCTGACGGAGCCACAGCAATGCTGCTCCAAAGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCACCAAAGTCCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCCTGCCAACCAAGTGTGCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAAACCTGCCCGAACCAAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGAGAAAG
AGAGGCCCGGGCACCCAGCCCCACTGGCCTCAGGCCCTCTGAGCTTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAACATAAGAAAGCCT
GTGTGCATGGCGGGAAAGACGTACTCCCACGGGAGGTGTGGCACCCGGCTTCCGTGCCCTCGGC
CCCTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGAAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGTC
CTCGCCACACATCGGTATCCCCAGCCCAGACAACCTGCGTCGCTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAACTTCCACTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTCCAGAAAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTGCAGTATGAGTGTATTATTATAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFS
SAHELFPSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQS
DEEDSVQSLHGVRHPQDPCSSDAGRKRGPGTPAPTGLSAPLSFIPRHFRPKGAGSTTV
KIVLKEKKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTC
PTEYPCRHPEKVAGKCKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDN
LRRFALEHEASDLVEIYLWKLVDDETEAQRGEVPGPRPHSQNLPLDSDQES
QEARNPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTCTCCTCTCTAATCCATCGTCACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGATCAGGGCAGTGGCAGGTGTTGGCCAGACAAGCCTGTCCAGGCCTGGTGGGGAG
GACGCAGCATTCTCCTGTTCCCTGCTTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTGAGGCTGGAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTCCCTAGTCTTACTACCAGAAGGCCATCTGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTCCCTCATTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGGAGATCTCTGACCCTCAAGAGAA
CGCCGGGAGCATACTGTTCCATGCGGATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTCGAGCCTATATCGTGGCACCTGGCTACAAAGTACTGGGAATACTCTGCTGTGGCTA
TTTTTGGCATTGTTGGACTGAAGATTTCTCTCAAATTCCAGTGGAAATCCAGGCGGAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCTGCGTTCTGATCTGAAAAGTGTAAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGAAACATTA
CTGGGAGGTGGACGGAGGGACACAATAAAAGGTGGCGGTGGAGTGTGCCGGATGATGTGGACAGGGAGGA
AGGAGTACGTGACTTTGTCTCCGATCATGGTACTGGTCTCAGACTGAATGGAGAACATTGTATTC
ACATTAATCCCCGTTTATCAGCGTCTCCCCAGGACCCACCTACAAAATAGGGTCTCCTGGACTA
TGAGTGTGGGACCATCTCCTCTTCAACATAATGACCAGTCCCTATTATACCTGACATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCCTCCCCAGGGTGAATGTAGGATGAATCACATCCCACAT
TCTTCTTAGGGATATTAAGGTCTCTCTCCAGATCAAAGTCCCGCAGCAGCCGGCAAGGTGGCTTCCA
GATGAAGGGGACTGGCTGTCCACATGGGAGTCAGGTGTATGGCTGCCCTGAGCTGGAGGGAAAG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGTATTTGAAATACCAACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCAATTACACTTCAGTAAAAAA

FIGURE 84

MALMLSLVSLLKLGSGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSSVH
LYRDGKDQPFMQMPOYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPGQDLSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSFFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRLNGEHLYFT
LNPRFISVFPPRTPKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGCCAGTAGTCATGGCTACTGGTCCGGAAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGCAGTGTGGAGGAGACTCGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCTGAGCATCAGAGATGCCAGAAGAAGTGTGCGGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCGAGGACCATGGCACCAGCCTCACCTGTCAAGGTGACCTTCCCTGGGCCAGCGTGACCA
ACAAGACCGTCCATCTAACGTGCTCACCGCCTCAGAACTTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCCAGAGGGCAGTCT
GCGCCTGGTCTGTGAGTTGATGCAGGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCTGTGCCCTCACAGCCTCAAACCCGGGGTGCTGGAGCTGCCTGGTGAC
CTGAGGGATGCAGCTGAATTACCTGCAAGAGCTCAGAACCCCTCTCGGCTCTCAGCAGGTCTAC
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTGGGGAGCTGGAG
CCACAGCCTGGTCTTCTGTCCCTGCATCTCGTTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGAGATA CGGCATAGAGGATGCAAACGCTGTCAAGGGT
AGCCTCTCAGGGCCCTGACTGAACCTGGCAGAACAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGAGGAGAGCTCCAGTATGCATCCCTCAGCTCCAGATGGT
AAGCCTTGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAAGACAGACAAATTCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFPGASVTNKTVHLNVSYPPQNLTMVFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSPNSPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGASQGPLTEPWAEDSPPDQPPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGC
CAACTACAACACCTTGGATCTGCAGAGGCAGCCACGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCTCCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGGATTCACTGCAGGGATTGTTCAAGTTCAGGGTATTTAATAACGAGAGAGCAGCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCTCCACCAGATCCAAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKS PMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGCAGGGAGACTTCAGGAGTCGCTGCTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCCCATGGGCCGGTCTCAGGGCTGTGCCCTCGCTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCAGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCCGCTCT
CTGTCACCCCTGGGCCTCTTGCAGCTGGAGCTGGCCGGTTCCCTCTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCA
ATTCGAGCGTTGGAGTGCACTACGTATTGGTACATTGGTCTTCTGCAGTGCCTCCAGCTG
TCACTGAAATGGCTTATTCGTACCGTCTTGGGCTGAAAAAGAAACCTCTGATTACCTTCA
TGACGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGTATTCTGGAAGAAGGAAG
GCATAGGCTCGGTTCCCTCGAAACTGCTCTGCTGGAGGATATGTGTTGAAATAATTACG
TCTTGAGTCTGGATTATCCGATTGTATTAGTGCTTGTAAATAAAATGTTGTTAGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGACGAGAGCGGAGGAAGATGCAACTGAC
TCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTCTATCTGGTATCTGTGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAACGCCCCTGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTGGCTGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTCATCGAAGCCAAGGCCCTCCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGGCCACCTGGAGCTGCTCCAGGCCCTCAAAGTCGTCTGTG
TCTACATGCCCTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
CATAGTGATAACCCCTACTACCCATCTGGGTGACCCGGGCCAGGCCACAGAGGCCAGGGC
TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGCTGGTCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGGCCACAGAGAGATGCTGG
TCCCCGAGGCCTGTGGCAGGCCATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAGC
CCTTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGAGATTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTCCTAGGAGCCAGTCAGCAGGGTGGGGTGG
GCCAGAGGAGCTCCAGCCCTGCCTAGTGGCGCCCTGAGGCCCTTGTGCTGTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATTCCCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACAGACAGGACGATTGTGGCTCCACACTAACGCCACAGCCCAC
CGCGTGCTGTGTCCCTCTCCACCCCAACCCCTGCTGGCTCCTCTGGAGCATTCCATGTCCCG
GAGAGGGGTCCCTAACAGTCAGCCTCACCTGTCAAGACCGGGTTCTCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGCGGGGGCCGGCAGAGCATGTGCTGGATCTGTC
TGTGTGTCTGTGTGGGTGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTGTTCTGGAGCAGGAAATAAGCTTCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRP MANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKVCVDYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCC**ATG**ACTGCGGCCGTGTTCTCGGCTGCGCCTTCATTGCCTCGGGCTGCGCTGCC
CTTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGGAGCGTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTGCATATTATAAACTCTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTGGCCTTGGAAATCATGA
GTGGAGTATTCCTTGTGAATACCTATCTGACTCCTGGGCCAGGCACAGTGGCATTCA
GGAGATTCTCCTCAATTCTCCTTATTCACTGACGCTGGTCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTAGCCCAGACCTTCATAAGTTCTTATTATGGAATAACCTGGCG
TCAGCATTATAATCCTGGTGCTCATGGCACCTGGCATTCTAGCTGCCAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTTGCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACGTGCCT
TTTCTGAAAATCCCTTTCTGGTGGATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISLWVFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGP GTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCAACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTGTCTTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTGGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTGTACACAACTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATTCCATTCTGTCCGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCCAACTCCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCAACTTTAAAAA
CAATAATTCAATGGATAAAATCTGTCTTGAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAACTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLGPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTPAGIQQRSTHAIIEATTEESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTGCCCAACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGCTGCTCCGCCCCAGGCTCTCCCCAAGGCCAGCCTGC
GTCCCTGCCTGGCTGTGGCTGCTTGTCTCCGCCCCAGGCTCTCCCCAAGGCCAGCCTGC
AGAGCTGTCTGTGGAAGTTCCAGAAAATGAGTGGAAATTCCCTTATAACCTGACCAAGTTGC
CGCTGCCCGTGAGGGGGCTGAAGGCCAGATCGTGTCAAGGGACTCAGGCAAGGCAACTGAG
GGCCCATTGCTATGGATCCAGATTCTGGCTCCTGCTGGTACCCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGTCCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCTTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGTACCAAGGCCTGGCATCCCCCTTCCCTGAGGCTTCAGACCG
GGATGAGCCAGGCACAGCCAACCTGGATCTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAAGTCTCCATCATAGAGAGCACCTGGGTGT
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCCTATACCCGCACCACATGGCCAGGTA
CACTGGAGTGGGGTGATGTGCACTATCACCTGGAGAGGCATCCCCCGGACCCTTGAAGTGAA
TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCATGGCAGGACTATGCCCTCCCCGTGACCCCACAGTCAGCATCCCTGAGCT
ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGACTAGACTGTCAGCAGAGGATGCAGATGCCCGGCTCCCCA
ATTCCCACGTTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGGAGAGCCTTC
CAGGTGGACCCACTCAGGCAGTGTGACGCTGGGGTGCTCCACTCCGAGCAGGCCAGAACAT
CCTGCTTCTGGTGTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTCAGCAGCACGTGTGAAG
TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTCCAGATTGGCCT
ATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCCATGCTAACAGCATTGATGCTGA
CCTCGAGCCGCCCTCCGCCTCATGGATTTGCCATTGAGAGGGGAGACACAGAACGGACTTTG
GCCTGGATTGGGAGCCAGACTCTGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGCCAGGCC
AGGCCCTGGAGCCACGCCACGGTACTGTGCTAGTGGAGAGAGTGTGCCACCCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTCCCCAGCCGGCTTTCTGCTGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTCTCCGGGAGGTGCACACCGCCCAGTCCCTGCAGGGCGCCAGCCTGGGG
ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTGCCCTGTGCCCTCCAA
TACCTCTGCACACCCGCCAAGACCATGGCTGATCGTGAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGCACGGTCCCTACAGCTTACCCCTGGTCCAACCCACGGTGCAACGGATTGGC
GCCTCCAGACTCTCAATGGTCCCATGCCTACCTCACCTGGCCCTGCATTGGTGGAGGCCACGT
GAACACATAATCCCCGTGGTGGTCAAGCCACAATGCCAGATGTGGCAGCTCTGGTGCAGT
CGTGTGTCGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGCCCGATGAAGGGCATGCCA
CGAAGCTGTCGGCAGTGGCATCCTGTAGGCACCCCTGGTAGCAATAGGAATCTCCTCATCCTC
ATTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
GAAGGCGACTGTCTGAATGGCCCAAGGCAGCTCTAGCTGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCCTGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGACAGAGTAGAAGCCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCACCAAGGATGCAGGCTGGACACCAACTT
TATGGACTGCCCATGGAGTGCTCCAAATGTCAGGGTGTGTTGCCAATAATAAGCCCCAGAGAA
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPPLEHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNI¹LLVLAMDLAGAEGGF²SSTCEVEVA³TDINDH⁴APEFITSQIG
PISL⁵PEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTEGT⁶FGLDWE⁷PD⁸SGHVRL⁹LC¹⁰KNLSY
EAAPSHEVVVVVQSVAKLVGPGPGATATVTVLVERVMPPP¹¹KLDQESYEASVPISAPAGS¹²FLLT
IQPSDPISRTLRFSLVNDSEGWL¹³CIEKFSGEVHTAQSLQGAQPGDTY¹⁴TVLVEAQDTALT¹⁵LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGP¹⁶NPTVQRDWRLQTLNGSHAYL¹⁷TLALHW¹⁸VEP
REHII¹⁹IPVVVSHNAQM²⁰WQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGT²¹LVAIGIFL²²I
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCTGGAAGCACCAGCCTTA
TCTCTCACCTCAAGTCCCCTTCTCAAGAACATCCTCTGTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTATGTTGGTCTACTATTGCAATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACGCCACCAACTCTGAGTCCAGCACACCCCTCCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCAGCAGCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCACACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTAGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTGTGACTGGA
ATGCACACAACCTCCATAGTGCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGCTCCAGTGG
GCCGTGGAAATCTCCTCATCACCTGGTCGGTTGTGGCGGCGTGGGCTTTGCTGGC
TCTTCTCTGTGAGAAACAGCCTGTCAGTGGAGAAACACCTTAAACACAGCTGTCTACCACCT
CATGGCCTCAACCAGGCCTGGTCCAGGCCCTGGAGGGAAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCTAACTGGGTCTGGAGGAGACCAAGTATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCTGAGCAGCCCCGGAAGCAAGTGCAGCATTCTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTCTTCATTCACTCCAGGAGACCCCTCCAGCTTGTGAGATCCTGAA
AATCTGAAGAAGGTATCCTCACCTTCTGCCTTACCAAGACACTGGAAAGAGAATACTATAT
TGCTCATTAGCTAAGAAATAATACATCTCATCTAACACACAGCACAAAGAGAAGCTGTGCTT
CCCCGGGGTGGGTATCTAGCTGAGATGAACTCAGTTAGGAGAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAACTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNMGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGLVPWEIFLITLVSVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPAGPGGNHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCGGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCCACGCACCCCACTCGGCGTCGCACGGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTGGAAATCAT
GGTGCATGGAAAGGGATTTACTTACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTACCCCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTGATAGGAAATGGAAGGATGACAAGAGCCATTCGAAGACATGATTGATTACTTTGTG
ATATTCAACGAACCACTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCAATCAGAGAACGACCTCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTCACGCCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCCTTCTATCAAG
GGGAGAAGAATTTTATTACCGGACAGAGTGTCAATTCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTCAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATTATAATCACCATTGTAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCTAACCTTCTAAGCTCAGATGCATTGCACTATGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGATGGGGTGAATAACTT
GGGCCAGAATATTATAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLGPGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFPREIHFHVRYPIDTLPTSKEQLQWLWCHKRWEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTGGTCTCAGTTCTACGAGCTGGTGTCAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCTGGTGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGGTCTTCAGGAATCAGTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCAGGGAGAACTGA
GTTTGAGAAGGACTCCATTGCAGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGTTCAAGTCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCAGTGGCTCACCTCCTCATTCCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCTGTCCCTCAGGCTGGTCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGAGCATATTGTGTTCCATCCACCTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCGCTGGCTTCTA
TTTACTCGGTTACTCTGTGGTGCCTGTGTGGTGTGTCATGGGATGATAATTGTTTCTC
AAATCCAAGGGAAAATCCAGGCGGAACGGACTGGAGAAGAACGACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACCGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTACCCGAAGCTGCG
TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGAAGAAC
ATGTGACTTTGTCTCCAACAATGGGTATTGGTCCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGTCTCCT
GGACTATGAGGGTGGGACCATCTCCTTCTCAATACAAATGACCAGTCCCTTATTATACCCCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTCATATGTCCAGTGTCTGGGAT**TGAGACAGAGAACACCCTGCTAAAGGGC**
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCGACAGGTGGCCAGCTCCT
CCGGAGCCTGCGCACAGAGACTCACGCCCTACTCTCCTTAGGGAGCTGAGGTCTCTGCC
TGAGCCCTGCAGCAGCGCAGTCACAGCTCCAGATGAGGGGGATTGGCCTGACCCGTGGAG
TCAGAACCATGGCTGCCCTGAAGTGGGACGGAATAGACTCACATTAGGTTAGTTGTAAAAA
CTCCATCCAGCTAACGATCTTGAACAAAGTCACAACCTCCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGAACAGAACATGTGAATCATGCTTCAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTATATTACATTCCCACCAAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACCAAATACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAGAGGGAGGTAGGATTTCACTGATTCTATAAGCCAGCAT
TACCTGATACCAAAACAGGAAAGAAAACAGAACAGAGAACAGGAGGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATATTAA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGAATTCAAGCACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFWQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCSSGWFPQPTAKWKGPQGQDLSSDSRANADGSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGVYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGTGGAACCCCTGGTTATCGGCCTCGTCATCTCATATCCCTGATTGTCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAACCTACAATTACTATAGCACATTGTCATTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTCTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAAATCTCCATTAAAGGAAGAATTGTCAAGTCTCAGGTTATCAAGTC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTGTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAAAAAATCAACAAGAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAAACCTAGGTCAGAGTCTCAGGATCGTGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGATGGAGTCATCGCTGTGGAGCAACCTTAATTATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTAACATATAAGAACCTGCCAGATGGACTGCTCCTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGTCTCCGGAGAATAATTGCCATGAAAATACAAACACCCATCACA
TGACTATGATATTCCTTGAGAGCTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGACAGGATTGGAGCACTGAAAT
GATGGTTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACTTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAAACAGATGCAT
GCCAGGGTGACTIONGGAGGACCCTGGTTAGTCAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTGTATAGAGTTACGGCCTGGGGACTG
GATTACTCAAAAATGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTGTTTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAATGTTGCTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCTAGAAGTTGTCAGAATTGACTGTTGACATAAATTGTAAT
GCATATATACAATTGAAGCACTCCTTCTCAGTCCTCAGCTCCTCTCATTCAAGCAAATATCCATT
TCAAGGTGCAGAACAAAGGAGTGAAGAAAATAAGAAGAAAAATCCCTACATTGACAGAA
AAGTATTAGGTGTTCTAGTGAATATTAGAAATGATCATATTCAAGGTCAAGCAAAGACA
GCAGAACATCAACTTCATCATTAGGAAGTATGGAACACTAAGTTAAGGAAGTCCAGAAAGAACCAAG
ATATATCCTTATTTCATTCCAAACAACACTATGATAAATGTGAAGAAGATTCTGTTGACCT
ATAATAATTACAAACTTCATGCAATGTACTGTTCTAACAAATTAAAGCAAATATTAAACATTG
TTACTGAGGATGTCAACATATAACAAATAAAATAAACACCCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPHSDYDISLAELSSPVPYTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:
amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACGCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCCGCGAAGAAGTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCCACTATCCCCAGGCAGGGCTGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTTGCCTGGGAGTAGGATGTGGTGAAAGGATGGGCTCTCCCT
ACGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCCTGCCTACGCCCTCAA
TCTGCTCTTGGTTAATGTCCATCAGTGTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACAGCAGGGTAGAGGAAGCAGTCATTTGACTTACTTCCT
GTGGTTCATCCGGTCATGATTGCTGTTGCTGTTCTTATCATGGTGGGATGTTAGGATATTG
TGGAACGGTAAAAGAAATCTGTTGCTCTTGCATGGTACTTGGAAAGTTGCTGTCATTTCT
GTGTAGAACTGGCTTGTGGCAGGGACATATGAACAGGAACATTGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTGGAATTTCAGAGAGAGTTAAGTGCTGTGGAGTAGTATATTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCCCTGCTGTGGTAGAGAATTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGGGAAGAAAATGTATTCCCT
TTTGAGAGGAACCAACAAACTGCAGGTGCTGAGGTTCTGGGAATCTCATTGGGTGACACAAA
TCCTGGCCATGATTCTCACCAATTACTCTGCTCTGGCTCTGTATTATGATAGAAGGGAGCCTGG
ACAGACCAAATGATGCCCTGAAGAATGACAACACTCAGCACCTGTCAATGCCCTCAGTAGAACT
GTTGAAACCAAGCCTGCAAGAATCTTGAACACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGGAGTTAAAAAGAAATGTCACAGAAGAAACCACAAACTGTTTATTGGACT
TGTGAATTTCAGTACATACTATGTTGAGAAATATGAGAAATAAAATGTTGCCATAAAA
TAACACCTAACATATACTATTCTATGCTTAAAATGAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCAACCAACTGTGTAGCC
TGTGTATGACTTACTGAACACAGTTGAGGAGCATGGTTGATTAGCATTCCGCA
TCCATGCAAACGAGTCACATATGGTGGACTGGAGCCATAGTAAAGGTTGATTACTCTACCAA
CTAGTATATAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATA
CTCAGCGATCTTCTGATGCTAAATAATTATATCAGAAAACCTTCAATATTGGTACT
ACCTAAATGTGATTTGCTGGTTACTAAAATATTCTTACCACTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTGGTATATAAGTCTGTGTTAAATCTGTATAATTCACTGATCATT
TTCACTGATAATTGAAATAACCATTATGAAAAGAAAATTGTCCTGTATAGCATCATT
ATTTTAGCCTTCCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTTATTAAACTAACACTAATTGAAAATTACAGTGTGATACATAGGAATCATTTC
AGAATGTAGTCTGGCTTTAGGAAGTATAAAAGAAAATTGCACATAACTTAGTGTGATT
AAGGACTTGTATGCTGTTCTCCAAATGAAGACTCTTGTGACACTAAACACTTTAAAAA
GCTTATCTTGCCTCTCAAACAGAACAGCAATAGTCTCAAGTCAATATAAATTCTACAGAAA
TAGTGTCTTTCTCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTGTTTACAA
GAGTATAGTATATTGAAATGGAAAAGTGCATTACTGTATTGTGTTATTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTGTCATAAAATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMWDPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTCTGATAAAGCCCTACCAAGTGCT
GATAAAGTCTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAGAGAAGGGACAAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTTCAGTGGCAGTGGTCCAGAAGGGATGCCTCCATTCTGCTTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGGCCCTAGATCCTGTTCTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTGGATGAGTCAGGTCCTCCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGTTGTGCCAGCT
TCAATGGGAACTGCTGTCTGGAACACCACGGTGGAAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCACGTCTACTGTGGTCATTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGTCAGATAACCAGCGAGTGCACATGCCTCCAGGAACACTGTGCTAGGCCCTGACAGGAGACATGCTTG
ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAAGATCTGTGTGAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCCGTGTGCTAAGAAGTGAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGCTGGTGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTTCCTGACCAACACCTCCTGCCAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCATGTGGTGAATGACAAGATTGTGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAAGCAAGCTGCTGATCCGGTG
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGATCTTCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTACC
GGGAAGCTCTGCCAACCTCAAGCTCGTACTCCCTACTTTGCATTGAGCCCTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTGCACCCCAACCTCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGGTAAAGCAGTACACATCCGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTTCAAGTTGTGGCAAAGACCACAAGGAAGTGTGTTCTGCACTGCCGGTTCTGTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGGGAATGCCGTGGGAGAGGACTCAGC
CGGTCTACAGGCCAGACGCTAACAGGCCGATCCGCATCGACTGGAGGACTAGTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCCTTTGGAGCTTCTCCCCCACCAGCCCTCAAGAACATCTGCCAACAGC
TGGGTCAGACTTCACACTGTGAGTCAGACTCCCAGCACCAACTCACTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGCTGAACAATGTGGCTGGGTGGGTTCATCTTCTAGGGTTGAAAACACTAAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTCTTCCACTAAATACCTCGTGTGGTCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAATAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAATTAGTTACCTTGAATTCAGTCAAATGCAGACTAA
TTATAGGAAATTGGAAGTGTATCAATAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCLWNTTVEVKACPGYYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNILFSLKTCGTVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRSLSYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCTGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGCAGTATCATCATTGT
GGTTGTCCTCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCTGGGGAGGACGAGGAGCAGTGTCAAGAGCTC
CCCAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACGGTTCTCTGCCTGTTGACAACACTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATCGGAACACTCAAGTGGCCCTGTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGTGGGAGGAGGCCTCTGGATTCTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATTGGACCCCCACTGGTCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCACCTGGCTGTGGCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCCACTCACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTCTT
GATGAGGAGCTACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGCTGACATACTGCTGCAGGCGTCAGTCCAGGTATTGACAGCACACGGTCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTCAGGCATCCGGAAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGCATCGTAGCTGGCTATGG
CTGCGGGGCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCACTGGAGCCCTTCCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTGCCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCCTCAATTCTGTAAGAGACCCCTCGCAGGCCAGAGGCCAGAGGAAGTCA
GCAGCCCTAGCTGCCACACTTGGTGCTCCAGCATCCAGGGAGAGACACAGCCACTGAACAAGGTCT
CAGGGTATTGCTAACCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCCAGCCCTGTCCGTCTCACCCATCCCCAA
GCCTACTAGAGCAAGAACCAAGTGTAAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCCTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIP III I ALLS LAS II IVVVL IKVILD KYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQM GYSRAVE IGP DQD LDV VEITEN SQELR MRN SSGP CLSGS LVSLHCLACG KSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVC GGSILD PHW VL TAAH C FRKHTD VFNWKVRAGSDKL
GSF PSLAVAK III IEFNPMYPKDNDIALMKLQFPLTFSGT VRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGD SGGPLMYQS
DQWHVVGIVSWG YGC GGPSTPGVYTKV SAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACGTCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTTTTGTTCTTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCAGTGTGGCTCT
GGCCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCTCGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCGCTGGAGCAGTGCCTCACCAACTGTCTACGTCT
GGAGGCAGTCACTCGGGCAGTAGCTGAGGCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCTGGCGTAGAAGGGATTGACAAGCCCAGATTCAAGGCGATGGCTCCACTGCC
AGGCATCAGCCTTGCTGTAGTCATCACTGCCCTGGGCCAGGACGGGCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTATGTCTGCACATCACCTG
ATCCATGGCTAATCTGAACTCTGCTCCAAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCCTGAAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTGTGCCAAATTATGGTCAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCATCAGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGCTGGAGGAGCAAGCAAAGTGACCATTCTCCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAAGGCCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGGCCCTTATACCCAGGAGACTTGATTGAATTGAAACCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCAGGGCGGGTAGATCACCTGAGGTCAAGGAGCTGAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTACTAAAAACAAAAACTAGCCAGGCATGGTGGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAACATCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAAATAAAAAAGAATT
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTAGCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGCTGGCCCTAACTCTAATTGTCTT
GTTTGGGGAGCAAGCACTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACTGG
CATCTACTTCGTGGTCTCAAAAATGTTTATCAAAACTCAGATTAAAGTATTCTGAATTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAACAGTCAGTG
ATTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAAAATTCCAAAATTCTGGA
GATTGTGATAACGTGACCATGTATTGGATCAATCCCCTCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGGAGGAGAAGATCTTCACTTCCTGCCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCCTGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTCGAGGCAACCGCTATTGCCGCCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGCTGTCTGCT
TAACTGGTGGTGGCCCGCATGCTGGGAGGGTCTAATAGGAGGTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTLIVLFWGSKHFPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTFEQSIVWVPAEKPIENRDFLKN SKILEICDNVTMYWINPTLISVSE
LQDFEEEEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTCACACCTCGGCAGCAGGAGGGCGGCAGCTCTCGCAGGCGCA
GGCGGGCGGCCAGGATCATGTCCACCACATGCCAAGTGGTGGCGTCCCTCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTCACCATCCTGGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCAGTGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCATTGTCTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACTCTGGATGTCCACAGCTAACATGTACACCGGATGGTGG
GATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGGCTCTGTTCTGGCTGGCTGGAGGCC
TCACACTAATTGGGGTGTGATGATGTGCATGCCTGCCGGGCTGGCACCAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCTCAGGCCACAGTGTGCCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGCAGTGTTAATGCTCTAACACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCAAAAAACAAGGAGATCCCACAGTGTGATGATCTTCTGCTTTGACTCACAGCTGGAGTTAGAAAAGCCT
CGATTTCATCTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTGTCTAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTCAATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATTTAGACAGACTCCCCCTC
TTCCCTCTAGTCAATAACCCATTGATGATCTATTCCCAGCTATCCCCAAGAAAACCTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTCCCCACCCCCACTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGCCCATGATCTGGTTTCTT
ACACTGTGATCTTAAAGTTACCAAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCTGAGCTCTCCACTGGAGTCCTCTTCTGT
CGCGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAAATTAAAGTCCTAAATATAGTTAA
AATAAAATAATGTTTAGTAAATGATAACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTGCATTGTCTATATGTTACTTGTAAAGTCATGCTTAAGTACAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTGGGAGGCTGAGGAGGAAGGATCACTGAGCCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATACACCTGTAGTCCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCCTGCACAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAATAATGGAACACAGCAAGTCCTAGGAAGTAGTTAAAACTAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGVWAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKIIDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAAACATCCCAGAGGATAAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTG
GTGGTGTGGAATGGTGGCACAGTGGCTGTCAGTGCATGCCTCAGTGGAGAGTGTGGCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGAAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTCCGTGATGTCCTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
CATCACGGGCATGGTGGTGCATCCCTGTGAGCTGGTTGCCATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTGTTCTGCTGCCTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATAACCTCCCATCGCACAAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTTAGTTGTATGTTTTAACTTTACTATAAGCCATGCAAATG
ACAAAAAATCTATATTACTTCTAAAATGGACCCAAAGAAACTTGTATTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTGATTCTATAAGCTATTGAGAATGAGATA
TTAACCCAAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGGTTCAAGCATTCTA
CTCTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAACTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTAAAACAGCTTAGGGATTAATGCTCCTCCA
TTTATAATGAAGATTAAAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTCTTATCCTCTTCCAGGGCTTTGCTTCAACTGCTT
CTTGTGTATTAAATTAAACATTAAAACGCAGATATTGTCAGGGCTTGCATTCAAACACTGCTT
TTCCAGGGCTACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGATGGTTAGGAAAGTG
AAAATTTTGTGTTGATTGAGTACAGACTTGAGGTTCATCAATATAAAAGAGCAGAAAATA
TGTCTGGTTTCAATTGCTTACCAAAAAACAAACAAAAAGTGTGCTTGCAGGGCTTGCATTCAAACACTGCTT
GCTCCTATGTGGGTACCTGAGTCAAATTGTCATTGTTCTGTGAAAATAATTCTTCAATTGTA
CCATTCTGTTAGTTACTAAAATCTGTAAACTGTATTGTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCACGTCTGTCTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRGTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCAGCGGGCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCTCCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCCTGCCCGAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCGCCTCCCCGAGCGGCTCCGCAGGCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCCGTCAGCGCCTCTGAGATCCCAAGGGAAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCAATGTTATTCCGGTACACCTGGATCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTGTTCACTGGCTCACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTTCAAGGACCTCT
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTACATCG
CACTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCT
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTCTACTGGATGGAATTCACTGGCT
ATTATTGAAGAACTACCAAAATAAATGCTTAATTTCATTGCTACCTCTTTTATTATGCC
TTGGAATGGTCACTTAAATGACATTAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTCACACTGTTAAATCTAGCATTATTCA
CTTCAATCAAAAGTGGTTCAATTGTTAGTTGGTTAGAATACTTCTCATAGTCACATT
CTCTCAACCTATAATTGGAATTGTTGTGGCTTTGTTCTTAGTATAGCATT
AAAAAATATAAAAGCTACCAATTGTAACATTGTAAGAATTGTTATATCT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAAASPQRLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGTACGGGGCTCTCCTGCCTCTGGGCTCCAACGCAGCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATAACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAACGTAAATATCCATGAAGATCC
TATTACCTAGGAAGATTGATGTTGCTGCGAATGCGGTGTTGGATTATTTGTTCTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCAAATCGGTCCATCTCCAAGGGTCCAATT
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGCTGTCATGCAACTG
GCCCTAAAGCAAAGCAAAGACCTAAGGACGACCTTGAACAATAACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTACTGACAATG
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTAAGGCAAATGGTATTGTGA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATACTGCTGGTTGCTTAGGTTGTCCTCGCT
ATAACAGCCTCAAAACTTAAGTATAATCAATTAAAGGGCTAACAGCTCACCTGGCTATAC
CTTGACCATAACCATAATCAGCAATATTGACGAAAATGCTTTAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTCTTAACAATACCTTCAGACCTGTGACAAATT
TACGAACTTGGATCTGCTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGCCGCAACCTGGAACCTTGGACCTGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGACTCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTCCAAGCTCAAC
CTGGCCCTTTCCAAGGTTGGTCAGCCTCAGAACCTTACTTGCAGTGGAAATAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTACAAAGGCTTGATTATCAGGAATGAGA
TCGAAGCTTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCTGAATCTGCAGCCCTAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAGCAGAAATATTGCTCCCTGTAAACTGGCTAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGT
ATCGATGCAGTGAAGAACTACAGCATCTGTGGAAAAGTACTACAGAGAGGTTGATCTGCCAG
GGCTCTCCCAAAGCCGACGTTAAGCCCAAGCTCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGAGCCACAGAGCCGGCCAGAGACCGATGCTGACGCCGAGCACATCT
TTCCATAAAATCATCGGGCAGCGTGGCCTTCCCTGTCGTGTCATCCTGCTGGTTAT
CTACGTGTATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAGCAGCGTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTATGTA
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGACCTGACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTCTAAAAGCT
GGGAAATAAGGGTCTTATTGAACTCTGGTACTATCAAGGAAACGCGATGCCCTCCCTCCCC
TTCCCTCTCCCTCTCACCTGGTGGCAAGATCCTCCTGTCCGTTAGTGCATTATAACT
GGTCATTTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAATCAATGTGAAGCTT
GAACCTCGGTTAATATAACCTATTGTATAAGACCCCTTACTGATTCCATTAAATGTCGCAATT
GTTTAAGATAAAACTCTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHMHISNIDENAFNGIRRLKELILSSNRISYFLNNTR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNLAFFPRLVSLQONLYLQWNKISVIGQTMSTWSSLQRSDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLFIGQEILDWSISLNDISLAGNIWECSRNICSIVN
WLKSFKGLRENTIICASPKELOQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFKHIIAGSVALFLSVLVLVLLVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCSYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTATCGTCTTGCCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCGCGCTACCGTGGCGAGCT
AGCAACCTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTATTGCCAGTT
TTAGCCAATCCAACGTACCTAGTGAAGGTTAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTACATCATGCATTGCAAAATCTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCAATATAACAAAGAGCAGCACTGGTGAATATGGGAGATT
ACCACTTATGATAACAGTGAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTTCTATTCTGGAACACCAGCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAAGGAAGGGACTTTGTATAATCATCG
ACTGACTGCTTGATTCAAGGCTGTTCAAGGTGAAGGATTGAGTCTATATAAAGGCTTTACC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLCVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASICTPADVIKSRI
MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGGCGGCGGCAGGAGAGCAGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTGGCTAGTCCAGCGAGGCGACGGCGGTGGGCCATGGCCAGGCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTACGGGGCCTGGGGGCATGGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCGCACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGACTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTGGCCCGGCCTGACACCCCTGCTCTCAGGCAGCACAGCAGTGGTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCAGGCCACCGAGTGTAA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCCAGGTGTGG
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTCCCTCCCCACCCCTCATGGCTTGCCCTGCCTCTGGATTAGGTGTTGATTCTGGAT
CACGGGATACCACTCCTGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGTGCCCTGGCACTTGGCCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCCAAATCCCCATCTTGACCTAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTATATCTGTGTTATCCAGGGCTCC
AGACTTCCTCTGCCTGCCCACTGCACCCCTCCCCCTATCTATCTCCTCTCGGCTCCCC
AGCCCAAGTCTGGCTTGTCCCTGGGTCACTCCCTCACTGACTCTGACTATGGCAG
CAGAACACCAAGGCCTGGCCAGTGGATTTCATGGTACATTAAAAAGAAAATCGCAACCAA
AAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININSMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTCTACATGGGCCTCCTGCTGCTGGTGCCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGCCCTCCCACCCCTCCTTGACTGCCGGCCGTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGCCCTGGCCCAGAATTCCAGTTCTGGTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTAAAGAGGGTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTCGCTGTTCAATTCAAGTGCACCCAGGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAATTCTCATGTATTTCATTTATTAA
TTAATTAACTGATAGTTGACATATTGGGGTACATGTGATATTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTATTTATTCTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCAAAGTAGCTGGACTACAGGCAT
GCACCACAATGCCCAACTAATTGTATTAGTAGAGACGGGTTTGCATGTTGCCAGG
CTGGCCTGAACTCCTGGCCTCAAACAATCCACTGCCTCGGCCTCCAAAGTGTATGATTACA
GGCGTGAGGCCACCGTGCCTGGCTAAACATTATCTTTCTTGTTGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGACTTCTCCCTCT
ATCTAACTGTATATTGTACCAAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTTACCTCCATGAGATCCACTTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGCTGGCTTACCTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTCGTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSPRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMPPFKRVFCQDGNVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAAACAAATAATGCTGGATTCCCTCAGAT
TTGAAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCAACACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAAGAATTACATTAGATGACAACACTCTGCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTCTGTCCTGCGTAATCACCTAGCACAATTCCCTGGGTT
GCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTCATCACCATCTTCAAG
GTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTGACAGAGCTGTCCTGGTGCAGATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATGGGTGCCCAATGCTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCATAATAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACTGATTCTCGAACATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCGTGGGATGGCTATTAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTCAAAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCAAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCAATCTCTGGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCCATAGCCGGATTGGATCTATAACAGAAACAATTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCTATAAAGTATGCATGGTC
CCATGGAAACCAGCAACCTCTACCTATTGATGAAACTCCTGTTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAACCCCAATTAC
TTTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTTACCATTGCCCTTCTGCTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGAAATCAGGGAAACTCTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGGAGTTGTAATACACACCATTTCCTCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSLTI PWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQTITAI PNTVYPAQGQWPAPVTKQPD
IKNPKLTQDQTTGSPSRKTITITVKS VTS DTIHISWKLALPM TALRLS WLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVC M VPMETSNLYLFDETPVC IETETAPLRM YNPTT LNREQEKE
PYKKNPLPLAAIIGGAVALVTI ALLALVCWYVHRNGSLFSRNCAYSKGRRKDDYAEAGTKKD NS
ILEIRETSFQMLPISNEPISKEE FVIHTIFPPNGMNLYKNNHSESSSNRSYR DSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561, 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCACTCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCCAGGCCACCCAGGCTTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGCCATGCCAGCCTGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGAGAAGTGGGCCAGGAGCCCCTGC
TGCAGCCGCTGAGCCTGCGGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAACAGAGGTTCCAGCTGACCCCTCCAGCCGG
GTTTGGAAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTGGTGTACCCACGTTGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTCCTGGTGCAGCTGGAAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTCTGGCCAGAATGAGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCAACCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAAATGGGCGGCTTCTCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGTTGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGGAGAAAAACAATTCCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTGCCCTGGTTCAAGC
AATTCTCTGCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGTGCACCACATGGCTAAT
TTTATTTTTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACCTCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAGGGCCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCTAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTCCTATACATCCTGGCAGAATAACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCC
TTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTTCAACTGCAACTGAACTGAAAAAA

FIGURE 134

MSARGRWEGGRRACRGLARAGAERVTSSEQRPMASLGLLLLLTLAPPLWSSLPGLD
TAESKATIADLILSALERATVFLERQLPENLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGDSSEPCGLSDLRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDI FMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGCTTCTTCCCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGTGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTCAGGGTGAAGC
CTCAGGACTGGATCTCGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCTTAAG
ACAGATGGGAGTTTGTGGTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTGATCCGTTGAGTGGATATCACTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTCTTACTTTATTAAAAGGAATCGTGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTTCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATTCAAACCATGAGTGCCTGAT
GTTCTGAGTCATGACAAGACTCTCTTCAAATCATCTGGCAAATCTAGCAGCGCAGCAG
TAAAACAGGCAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTTGGAAACCGTGTGAAGCAACTACTATAACTT
GAGTCATCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTAGCACATGTTGTACTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACTATTATGTATATTAAATTAAAACATCTTAATCCAGAAATCAAAAAAAA
AAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLLTDGSFVVHDIPSGSYVVEVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCTCCCCAGTTCCCTGTTGGCTGAGGGGA
CCAGAAGGGTGGCTACGGCTTCTGGAAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGAGCCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTTCACGGGAGGCTGGCAGT
TTTCTTACTCCTGTGGTCTCCAGATTCAAGGCTAAGATGAAAGCCTCTAGTCTGCCTTCAGC
CTTCTCTGCTGCCTTATCTCCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTTCAGGAAATACGAAATGGATTTCTGAGATAACGGGCAGTG
TGCAAGCCAAGATGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAAGACCCCTGACCATTATACTCTCCGGAAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTGGGGGAACTAGACATTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGCATGCCAA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGCACAGTGTATCTTATTATGCATTACTG
CTTCCTGCATGATTGCTTATGCATCCCAATCTTAATTGAGACCATACTGTATAAGATT
TGTAATATCTTCTGCTATTGGATATTTATTAGTTAATATATTATTATTGCTATT
ATGTATTATTTTTACTGGACATGAAACTTAAAAAAATTCACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAACTAAGGACATATTACTCATGCTGATGCTGTGAGAT
ATTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCACCCAGTAGACTCCCCAGTCCATAATTGTGTATCTTCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTTGCAACCAAAAAAA
AAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTCCGCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTCCCTCGGGATTCTCCCGGCTCCGCTGGAGCCAGTCTAAGTGGACACGCTGCC
CACGGAGCGGAGCCCCAGCGCCGAACCTCGGCTGGAGCCAGTCTAAGTGGACACGCTGCC
ACCACCTCTTCAGTAAAGTTATTGTTCTGATAGATGCCTTGAGAGATGATTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTGTTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTGTCACGTCACTCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATT
TTCCCAAAGCATTGTGGAATATGATGGAACAACTCATTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTGGATAAAAGTATTAAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGCTGGACCACATTGCCACATTAGGGCCACACTGAGGAGGAAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCTTACCCATTGCTGGTTCTTGTTGACCATGGCATGTCGAAACAGGAAGTCACG
GGGCTCCTCACCGAGGAGGTGAATACACCTCTGATTAAATCAGTTCTGCGTTGAAAGGAAA
CCCAGTGAATCCGACATCCAAGCAGCTCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTGCAGCTAGTAAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAAGATCCTGGGTTGAGCAGTTAAATGTCAGAAAGATTGCATGG
GAACGGATCAGACTGTACTTGGAGGAAAGCATTCAAGTCCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTGCTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCA
CTGTCATCTCCTGGTTCTGCTCTTATTGGTGAACAGCTGGCTTCTTCGGCCGTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTGCTACTTCTGAGCTGCTTGTGGCCTCTCGTGGCTGGCGCAGGCT
GCCTTCGTTTACAGACTCTGGTTGAACACCTGGTGTGCAAGTGCAGTGGCAGTGCCCTGGAC
AGGGGGCCTCAGGGAAAGGACGTTGGAGCAGCCTTATCCAGGCTCTGGGTGCCCCACACAGGTG
TTCACATCTGTGCTGTCAGGTCAAGTGCCTCAGTTGGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCCAAGGCGACGTTCTGTTGGAATTCTTAGTCCTGGCC
TCGGACACCTCATTGTTAGCTGGGAGTGGTGGTGGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGCCCC
ACCCCAACCCCTGCACAGCCCTCATCCCTCTGGCTTGAGGCCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTGTCATCAGGGCACAGGCCCTCTGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGCTCATGCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTTACAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDLILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTCCAGGTTATCGTACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGGAGTACA
TCCCTATAGAAAACAACGCCAGCACCTTAAGACCACTCACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCA
GCATTGACCA
GAGGATCACAAAGTACTGGCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAA
ACTACATACGCCAGAGATCTTCTTGCA
TTAGCCTCATCCTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAGTTGTCT
TAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTCAGCTGAAGAAGGAGAA
ACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGGCC
TTCATCTTTATAGGGCTCAGGTGGCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCGGATGGTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGTGACAGATAAATTGAGAACAGGAAACACATTGAATTTCATTCAACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGT
CAGCGATTAGGAACTGCC
CATTGAACGCC
TTCC
TCGCTA
ATTGAACTAATTGTATAAAA
ACACCAA
ACCTGCTCACT

FIGURE 142

MLLLLEYNFPIENNQHLKTTHTFRVKNLNPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGA
TCCAGGATCCTGCTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGCCTGCCGTGG
AAGGGAGGTCTGCTGGCGCTGCTGCTGCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCTGGAAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAAACTTTCACCTGCTTCTCACCACAGCACCAAGGCCCTGGATGACTCAGTCAGC
CTCCTGAACAAAGACCTGCTGGAGGGATTCCACTTGAGTGAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTAGAAATCC
ACACATCTTGAGCCTAATCATGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGGATGAGGAAG
CACCTGAGCTGGTGGCTGGCCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTACCGGCTGCATCAATGCCA
CCCAGGCGCGAACCAAGGGGAGTTCCAGAACAGACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTTGCTCCCTCAAGCATTGCGAGTTGGAGAGGGCGCAGGACT
TCGGGTCAACATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCCACGTATGCCCTGGTATGTGCCTGCCTGATAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCCACCTGCTCACAGTACTTCCAACAACTCTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCAGC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTGTCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTATCACTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACCATGTCGGTCCAGAGTCTCATT
CCTGATGATTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTGGCCTCCTCT
CTGTCTTCTTCCCTTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATTTCACTTGCTTGTCAAGGGTAGGTCACTGAGCTTAGTTTATTTTGAAATT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTTACCTGGAGACTTGAC
TCCCGCGGCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCAACACCCACCCACCCCTGGCTCTTCCTGTTTACTCCTCCTTCAATTATA
ACAAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACCTGGATTCTGGTTAGTGCTCCGATTCAAGCTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTGAATGAACAGATGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGACTCCTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGGCCGTGTTGACAAGATGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTACAAAAATTAACTCAAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAAGGAGAACTAAAACCTACAGTGAAGACAACTTGAGGAACCTCA
ATATTCCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGAAGGAAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTCTACCTTGAAACTGGATGAAATGATTGCTCTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAATATTGAATGGTGAAGAAACATGACAAAAAGGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAACAAAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAAGGAAAGATCCAGGAGTCTTCAA
CTGTTCTGAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAATTGGACCCAAAGG
GTTATTAGAAAGTGCTGAATTACAGTAGTTAACCTTTACAAGTGGTAAAACATAGCTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTGTATGTAAGCTGAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKTPMAAIQDGLAKGENDETVSNTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGCCTGAGTGGGCGCTGTGCTTCGAATGAAGGACTCGGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCAATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTATCCTGGGTGTCAGGGTGGAAAG
CCAGTGCCTGTCATGTGGGTGGGCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCAAGGAATCCAAGAGCTTACCTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAACTCACCCAGCTCCCGAGAATGGTGGCTGGAATGCCCATCACAGACTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACACTCCCTGGCAGAGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCCTCCCCACTGGATGGTCTACTGCTGTGGAACCTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGAGTGGTGGGAATCTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCCGAGGCAACCCATCCCCAGTTGAGCCTTATA
GGTCAGTAGCTCTCCACATGAAGTCTGCACTCACCACGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGCCCTGGCCCAGCCCCACCCCTTAACTCTGCCACTGTCATA
TGCTACCTTCCTATCTCCCTCATCATCTTGTGGCATGAGGAGGTGGTGTAGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTCTATTCCATGAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAAGGGCATATTGCAATTATTAAATTAAAGATAACCTATTATATT
TCTTATAGAAAAAGTCTGGAAGAGTTACTTCAATTGTAGCAATGTCAGGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTCCTAATTTCCTACAATGAAGATGA
ATTCTGTATAAAAATAAGAAAAGAAATTAAATCTTGAGGTAAGCAGAGCAGACATCATCTGA
TTGTCCTCAGCCTCCACTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTGGTGG
TTGTAGTAGTGTAGCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTGT
GTGGCTGGAATCTCTGGTAAGGAACCTAAAGAACAAAAATCATCTGGAATTCTTCTAGAAG
GATCACAGCCCCGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGACTGGTGA
ATTGTGTCCTCTCAAATTCACATCCTCTGGAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAGCCACAAGCTAAGAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAAGAGGCAAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAATTGGCTGTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCTCGTGTACATTCTGTGTGTCCCCTCCACAATGTACCAAAGTTGTCTTGTGACCAA
TAGAATATGGCAGAAGTGTAGGCATGCCACTTCAAGATTAGGTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
GCCACAGGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTGTT
AAGTTGCTCAGTTGGCTAACCTGTTATGCAGCAATAGATAAAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEEISVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCTGGCCACCAGCTGCCTCCTCTTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCGCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCTCCTGCCAGGCTCAGAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCAAATGAACCCCTGCCTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAACGATAGATATTGATAACATTCACTGTAACGGTGTTC
TATACACAGAAAACAATTATTAAATAATTGTCTTCCATAAAAAGATTACTTCCAT
TCCTTAGGGAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATAAGACTGCATTTATTATCATTATTAAATATGGATTATT
AGAAACATCATTGATATTGCTACTGAGTGTAAGGCTAATATTGATATTGACAATAATT
AGAGCTATAACATGTTATTGACCTCAATAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTIATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFHGVMSERCYLMQVLNFTLEEVLFQPSDFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTCCACGAGGCCTGT
CAGTCAGTGCCGACTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTAGTCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGTGAGAAACTGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCAGTGCCTGTGCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGCCAGTGAAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCAACGCCGT
TGCCTGTGCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAAACAGACTGTCTTACAGGCCATGCCATGGCGAGAAGGGACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCAGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGGAGCAG
CAGGATCCCAGGACAGGATGGGGCTTGGGAAAACCTGCACCTCTGCACATTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCACTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTTCCCTTTCCATCCCTGCTACCTG
GCCAGCACAGGCACCTTAGATATTCCCCCTGCTGGAGAAGAAAGAGCCCTGGTTTATT
TGTTGTTACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTTTATTGAATGTATAGAGATTATCAAATAAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATG**TCGCTCGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTGTGTGACGGGAAAAGCAACTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTCTGGTGGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTCATTGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAAGCCTGTGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACACCCTCCCCTGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCCTACTTGTGGC
AGCGACTGCATCCGACATAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTCCCT
GGATAACAAACAAAGCAAGCCGGGAGGCTGGCTGCCCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTT
TCTACCACCAACTACTGCCCTTCAATTAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAAATTGTTACTTCACTGAATTCTCAAAACCATTGAGAAGTGGTCACTGCTTGGCCA
AGTGGCAGAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTGCCACTAAAAGAAGGCA
GCAGACAAAGTCGTCTCCTCTTCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTTTCCCTGCCTTAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGTACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCCCTTGT**AG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS梧
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNK
SKPGGWLPLLLSSLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFFEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCAGCTCGGAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCCTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTGGGACCCCAACCGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLSILGLAFLSEAAARKIPKVGHFFQKPECPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTGGAGGCTGAAGAGGGATTC
CAGCCCCCTGCCACCCACAGACACAGGGCTGACTGGGGTGTCTGCCCTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTGCT
TGGCACTGGGCCGAAGCCCAGTGGCCTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCCGGGCCTCTGCCCTCTGGACAGTGACATACTCTGCCTGCCTGGGACAT
CGTGCCTGCTCCGGGCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGCCACTGGCGTGCATGGGACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGTGGAGGAGCCTAGGAA
TGCCTCTCCAGGCCAAGTCGTGCTCCTCCAGGCCAACCTACTGCCGCTCGCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTGTGCAGTTGGTCAGTCTGTGGCTCTGTGGTATATGAC
TGCTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGCCTATACTCAGGCCAGGTACGAGAA
GGAACACTAACACACACAGCAGCTGCCCTGCCCTGGCTAACGTGTACGCAGATGGTACA
ACGTGCATCTGGTTCTGAATGTCTTGAGGAGCAGCACTCGGCCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCAAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTGAA
CCACACAGACCTGGTCCCTGCCTCTGTATTCAAGGTGTGGCCTTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGCAGAGCTGGCTGCTGGACGCACCGTGCTGCCCGCAGAACGGCA
GTGCTGGCGGGCTCCGGTGGGGACCCCTGCCAGCCACTGGTCCACCGCTTCCCTGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTTCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTAGGTG
AACAGCTGGAGAAGCTGCAGTCAGGAGTGCTGTGGCTGACTCCCTGGGCCTCTCAAAGA
CGATGTGCTACTGTTGGAGACAGGAGGCCCCAGGACAACAGATCCCTGTGCCTGGAACCCA
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCAGGGCAGCTGCCCTGGCTGAGGCTTGTGAA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGCCTG
CCCCATGGACAAATACATCCACAAGCGCTGGCCCTCGTGTGGCTGGCCTGCTACTTTGCCG
CTCGCTTCCCTCATCCTCCTCTCAAAAGGATCACCGAAAGGGTGGCTGAGGCTTGTGAA
CAGGACGTCCGCTGGGGCGGCCAGGGCCGCGCCTGCTCTACTCAGCCGATGA
CTCGGTTTCAGCGCCTGGTGGCGCCCTGGCGTGGCCCTGTGCCAGCTGCCGTCGCGTGG
CCGTAGACCTGTGGAGGCCGTGTAAGTGAGCGCGCAGGGGCCGTGGCTTGTGAA
CGGCCAGACCCCTGCAGGAGGGCGCGTGGCTTGTCTCTCCCGTGCAGGCC
GTGCAGCGAGTGGCTACAGGATGGGTGTCCGGGCCGGCGCACGGCCCGACGACGCC
GCGCCTCGCTCAGCTGCCTGCCGACTTCTGCAGGGCCGGCGCCGGCAGCTACGTGGGG
GCCTGCTTCGACAGGCTGCTCCACCCGGACGCCGTACCCGCCCTTCCGACCGTGCC
CACACTGCCCTCCCAACTGCCAGACTTCCCTGGGGCCCTGCAGCAGCCTCGGCC
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCGGCCCTCAGCCAGGCC
CATCCCCCGGGACTCCCGCGCCGGACGCCGGGACAGGGGCC
CGGGACTTAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSCRWDSDLCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVLLEV
QVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSDGDNVHLVLNVS
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLLTQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPLSWENVTVDKVEFPLLKGHPNLCVQ
VNSSEKLQLQECLWADSLGPLKDDVLLLETGPQDNRSRLCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATTTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTGA
AAACATCCTGACGTGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTGAGCAGGATCACCCGGAAAGTCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTACCGCT
GTCAGTGCAGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATCGATTAGATGTTCATCCTA
CCCCCACGCCAATCCGTGCAGGCAGGCCACCAGCTAACCCCTGGAAGACATCTCCATGACCTG
TTCTACCACTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCTGACCCCTGACACAGAGTTCTGGCACCACATGATTGCGTTCCCA
CCTGGGCAAGGAGAGTGCCCCCTACATGTGCCAGTGAAAGACACTGCCAGACGGACATGGACC
TACTCCTCTCCGGAGCCTTCTGTTCTCATGGGCTCCTCGTCAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGCCGCTGCACCTCCAACTCCCTGAACGTCCAGCAGTCCTGACTT
TCCAGCCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTGACCTCAGCGGCCCCAGC
AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGGCCAGGAGC
TCCACAGCGGCATAGCCTGTCCAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCATCCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCATTCTACGCCCCACA
GCCATCTCTAAGGTCCAGCCTCCTCATGCCCTCAAGCCACTCCGGACAGCTGGCCTCC
CCTATGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTCTAGTCCTAAA
CACCTAGGCCTAAAGTCAGCTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCCTGGCTATGGAGGAATCCAAGAAGCAAATCTGCACCAGC
CCCTGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGCCAGCTCCCCCTCCTCTCAGTCCAGATCGAGGGCACCCATGTC
CCTCCCTTGCACCTCCTCCGGCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCTTGTGTCCAGGATGAAGCCAAGAGCCCAGGCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACCTGGATTCTCTTTCAGAGGCCTGGCCCTGACTGTGAGTGGAGTCGT
AGGGGAATGGGAAAGGCTGGTGCCTCCCTGTCCCTACCCAGTGTACATCCTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGTGCCTTGAGAGAAC
AGAGGGAGTGCATGCAGGCCCTGCCATGGTGCGCTCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTGGGGAGCAGCTGTGAGACAAGCGCGTGCCTGAGCCCTG
CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAAACTCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCTCAGGGATTGCTCTTGGCCCTGGGCTGAGACCTGGGAGTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCCTTGTGAGTGTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCAAGGCCCTGGAAAGAACAGAGGAGGCTGGCAGAACCAACACTGC
ACTTCTGCCAAGGCCAGGGCAGCAGGACGGCAGGACTCTAGGGAGGGTGTGGCCTGCAGCTCA
TTCCCAAGGCCAGGGCAACTGCCTGACGTTGCACGATTCACTTCAGCTTCACTCTGTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTCTGCAGGCAGGAGTTTCAGACCT
ATCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAACACAC
TGTACTGATGTCACAACCTTGCAAGCTCTGCCTGGGTTCACTGCAGGCCATCTGGGCTCAAATTCCAGC
CTCACCACTCACAAGCTGTGACTTCAAACAAATGAAATCAGTGCCTCAGAACCTCGGTTCC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTAAGATGAAATGAAGTCATG
TCTTAAAGTGCTTAATAGTGCCCTGGTACATGGCAGTGCCAATAACGGTAGCTATTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTGVSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSRAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESVCPKDEAKSPAPETSDEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGTGGCCACAACATGG
CTGCGGCGCCGGGCTGCTCTCTGGCTGTCGTGGCGCTCTGGTGGTCCCAGGCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGGACCTCAAAGTGTGCAGGGACGAAGAGTCAGCAT
GTTAATGTACCGTGGAAAGCTCTGAAGACTTCACGGGCCCTGATTGTCGTTGTGAATTTA
AAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAACCTTGGCTGGA
AGTGTGAACACAGTTGGATATTCAAAAGATTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACACTTGTCTGCTTGAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGCGTGA
ACTTGACCTGTGCCTGAGCCCGAGGCATTCAAGCTGATTCAAGGGATGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGGCCACCCACACCAGCGGT
CCTCGGGCTAACGCTCAGGGAGTGCAGTCTCGTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGAAAGCAGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCAGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATAAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTGAGCCAAACAAAAATATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVC GDEEC SMLMYRGKAEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADSEDEGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAACCTCCAGCCCCCACATCCACGCGGGCACAGG
CGCGGCAGGCAGGTCCCGGCCGAAGGCATGCGCGCAGGGGTCGGCAGCTGGCTCGGG
GGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAAGACTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGGATGAGCCCGTGG
TCTCGCTGCTGCTGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GCCAAAAGGTGTGTTGCTGACTTCAGCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTCAGGAGGCACGCCTGGCTGTGAGAGTGAGGGAGGAGTCCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAACAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTCTGATGGTGATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCTGCGGAAGTAAAAGTGTGTTGATGTATCACCAACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTACCAAGTGGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTTGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTTGCTTTGGAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAACTAGTCAAACCAAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTTAATAACTCATTGACTGGTTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTGGAAATGGCTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTATGTCTATTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCCAA
ACTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTGGGAGTA
TGTGTGTTAGAAGCAATTCTTTATTCTTACCTTCACCTTCATAAGTTGTTATCTAGTCATGTA
TGTATATTGTATTGAAATTACAGTGTGCAAAGTATTACCTTGCATAAGTGTGATAAAA
ATGAACTGTTCTAATATTATTGATGGCATCTCATTTCATGCTTGTGATAAGGATTTGATTAAAG
AAACTTATTACTGTTGTCACTGAATTCACACACACAAATATAGTACCATAGAAAAAGTTGT
TTCTCGAAATAATTCATCTTCAGCTCTGCTTTGGCAATGTCTAGGAAATCTCTCAGA
AATAAGAAGCTATTCTATTGATGAAAGTGTGATAAACACCTCCTCAAACATTACTTAGAGGCAAGGAT
TGTCTAATTCAATTGTGCAAGACATGTGCCTATAATTATTTAGCTAAAATTAAACAGATT
TTGTAATAATGTAACCTTGTAAATAGGTGCATAAACACTAATGCAGTCATTGAAACAAAAGAAG
TGACATACACAATATAATCATATGTCTTCACACGTTGCCTATAATGAGAAGCAGCTCTGA
GGGTTCTGAAATCAATGTGGCCCTCTCTGCCCCACTAAACAAAGATGGTTGTCGGGGTTGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATAAAAGAGGTATGTGGTTGAGACCAGGTGAATAGTCACATCAGTGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGAACATCGAAATCGTGTGAAAATGGGTTGG
AACCCATCAGTGATCGCATATTGATGAGGGTTGCTTGAGATAGAAAATGGTGGCTCCTT
CTGCTTATCTCTAGTTCTCAATGCTTACGCCTGTTCTCAAGAGAAAGTTGTAACCT
CTGGTCTTCATATGTCCCTGTGCTCCCTTAACCAAATAAAGAGTTCTGTTCTGGGGAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSSLGALLCGHGFACRRVSGQKVCADFKHPCYKMAFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217